MASTER IN COMPUTER SCIENCE

A Software Methodology applying Privacy by Design Principles for Mobile Hybrid Development

Barette, J

Award date:
2012

Awarding institution:
University of Namur

Facultés Universitaires Notre-Dame de la Paix, Namur
Faculté d'Informatique
Année académique 2011–2012

A Software Methodology applying Privacy by Design Principles for Mobile Hybrid Development

Jonathan Barette

Mémoire présenté en vue de l'obtention du grade de
Master en Sciences Informatiques.

Abstract

The world of technology has evolved in recent years giving way to smaller devices. Current, mobile phone are many on the market allowing the user to connect anywhere on the Internet. The consumer can take advantage of many services offered for fun or work. To satisfy this need, many companies develop their own applications and offer a more attractive access to information. But the competition is ubiquitous, the appearance often determines whether the user loves or not the mobile application. These custom applications often use personal data to provide customized functionalities. To the detriment of the user who has no idea that his information can be used without his prior consent. This manipulation may therefore have an impact on privacy. Existing since several years, this phenomenon is not close to stop. To avoid this, competent authorities try to include the right to data protection into development processes by providing appropriate solutions. Concepts such as Privacy By Design, PIA, PET,... are only part. Even associations that wish to improve the mobile eco-system propose privacy guidelines to encourage developers to take into account the data protection in their application. This scientific work aims to allow companies to integrate these principles into their existing mobile development methods without having a significant negative impact on productivity. However, we limited to provide a solution for certain types of applications.

Keywords: Privacy By Design, Data protection, Mobile application, Software Development, Agile, Proactive approach.

Résumé

Le monde des nouvelles technologies n'a cessé d'évoluer au cours de ces dernières années laissant place à des appareils de plus en plus petit. Les smartphones se font aujourd'hui la part belle sur le marché permettant à l'utilisateur de se connecter partout sur internet. Il peut ainsi profiter des nombreux services qu'ils lui sont proposés pour se divertir ou travailler. Pour assouvir ce besoin, de nombreuses entreprises développent leur propres applications proposant ainsi un moyen plus attrayant d'accéder à l'information. Mais la concurrence fait rage et c'est celle qui proposera la plus belle visuellement qui gagnera la compétition. Ces applications utilisent souvent des données à caractère personnel pour proposer une solution personnalisée. Ce au détriment de l'utilisateur qui n'a aucune idée que ses informations peuvent être manipulées sans son accord pouvant ainsi avoir un impact sur sa vie privée. Existant depuis plusieurs années, ce phénomène n'est pas encore réellement près à s'arrêter Les autorités compétentes essayent tant bien que mal de faire promouvoir les droits des utilisateurs en proposant des solutions adéquates. Des concepts comme le Privacy By Design, EFVP, PET ou autre n'en sont qu'une partie. Même les associations pour l'amélioration de l'éco système des applications mobiles mettent au point des recommendations pour insciter les développeurs à prendre en compte la vie privée lors du développement. Ce travail a pour but de permettre aux sociétés d'intégrer ces principes dans leur méthodes de développement mobiles actuelles sans pour autant avoir un impact négatif important sur leur productivité. On s'est cependant restreint à fournir une solution adaptée à certaines catégories d'applications.

Keywords : Privacy By Design, Protection des données, Application mobile, Développement logiciel, Agile, Approche proactive.

Acknowledgements

It would not have been possible to write this master thesis without the help and support of the kind people around me, I would like to acknowledge:

- My supervisor, Mr. Petit Michael for the support he gave me offering invaluable assistance and guidance. He was always accessible and willing to help his students with their research.
- Mr. Colin Jean-Noël, my professor of security, who has given me great suggestions during his lessons about privacy consideration.
- The faculty of computer science for giving me a variety of skills that helped me to the drafting of this scientific work.
- And, finally, my family and all my graduate friends for their support throughout the writing.

Contents

Chapter 1

Introduction

1.1 Context

In a world where technology is constantly evolving, a new trend has emerged, the mobile internet. It is now possible to surf the internet to make transactions of daily life using mobile devices. But are there risks to use this new type of revolutionary means for people?

1.1.1 Pervasive Computing

At present, the interaction between man and machine is at its peak. We see more and more technological devices, of daily life, which are interconnected and share information. The omnipresence of computing in our environment has lead to the use of the expression "Pervasive Computing" that *"describes the kind of computing that will result from the convergence of three computing-communication trends"* [1].

Ubiquitous Access to information has become commonplace in the civilian world that we know, especially through the Internet. More and more devices became networked, starting with the first computers then technologies providing the means to access information such as personal digital assistants, mobile phones or other, whether through wireless or terrestrial networks.

Embedded The computer is not limited to workstations, it also covers all elements miniaturized and integrated in household products, medical, industrial,... *"Capable of wireless communications, these devices may even be smart enough to self-organize into localized networks."* The embedded technology called RFID[1] tags is used to retrieve data remotely using appropriate readers. These tags are smaller and less expensive to produce and are mainly used to replace universal barcode. Embedded elements are ubiquitous in everyday life.

Animated Most current systems require information coming from external part of the physical world around us. These are measured by sensor instruments and can operate autonomously without human intervention.

[1]RFID : Radio Frequency IDentification

The "Pervasive Computing" suggests the influence of IT on our daily life and poses new social, cultural and psychological research challenges. Today, **mobile computing**, a particular technology contributing to ubiquitous computing, has a definite impact on the information sharing.

1.1.2 Mobile Computing

In constant evolution, **mobile computing** has become part of the daily live. It is *"a generic term used to refer to a variety of devices that allow people to access data and information from where ever they are"* [2]. Several categories of portable devices can be showed:

- **Portable computer** that can be easily moved from place to place. It is also called "Transportable PC".
- **Tablet computer**, often without keyboard, is shaped like a slate or a paper notebook.
- **Personal Digital Assistant** is a small computer with limited functionality.
- **Smartphone** is a mobile phone with a range of features and installable applications.
- ...

A smartphone is primarily a personal object for communication, but recently it became an indispensable tool for realizing everyday operations (buying online, chat with friends, send e-mails,...) without necessarily being at home. Today, most people have a smartphone, not only for calling but mainly for accessing to information via applications. This tool becomes more democratic with time. It offers more services to any person to make life easier. On the market, there are several types of mobile applications, which are free or not. Gartner, a mobile market research company, identifies the top ten of consumer mobile applications for 2012[3]:

1. **Money Transfer** - This service allows people to send money to others using Short Message Service (SMS). Its lower costs, faster speed and convenience compared with traditional transfer services have strong appeal to users in developing markets, and most services signed up several million users within their first year.
2. **Location-Based Services** - A part of context-aware services that will be one of the most disruptive in the next few years. It includes services to identify a location of a person or object and is used today in Social Networking as an entertainment service.
3. **Mobile Search** - The ultimate purpose of mobile search is to drive sales and marketing opportunities on the mobile phone. It is an evolving branch of information retrieval services that is centred around the convergence of mobile platforms and mobile phones.
4. **Mobile Browsing** - With cellular network, this type of application allows to display Web content and is specific to small devices. *"Mobile browsing is a widely available technology present on more than 60 percent of handsets shipped in 2009, a percentage Gartner expects to rise to approximately 80 percent in 2013."*

[2] Definition of mobile computing into glossary of `http://operationstech.about.com/od/glossary`
[3] `http://www.gartner.com/it/page.jsp?id=1230413`

5. **Mobile Health Monitoring** - It is the use of IT and mobile telecommunications to monitor patients remotely, and could help governments, care de-livery organizations and healthcare payers reduce costs related to chronic diseases and improve the quality of life of their patients.

6. **Mobile Payment** - Refer to payment services operated under financial regulation and performed from or via a mobile device. It is also an extension of online payment for easy access and convenience and an additional factor of authentication for enhanced security.

7. **Near Field Communication Services** - This new technology allows data transfer between compatible devices by placing them close to each other, within some centimeters. The technology can be used, for example, for retail purchases, transportation, personal identification and loyalty cards.

8. **Mobile Advertising** - A form of advertising via mobile phones or other mobile devices. It will be an important way to monetize content on the mobile Internet, offering free applications and services to end users.

9. **Mobile Instant Messaging** - A presence enabled messaging service that aims to transpose the Internet desktop messaging such as ICQ or MSN experience to the usage scenario of being connected via a mobile/cellular device. Mobile IM presents an opportunity for mobile advertising and social networking, which have been built into some of the more advanced mobile IM clients.

10. **Mobile Music** - Digital music service facilitating the purchase and download of music via mobile phone. Spotify, MOG, Rhapsody, and the iTunes Store are all examples of services that offer music downloads to mobiles.

These applications are easily available on the web to provide the user a range of services relating to everyday life. But, in most cases, they use consumer's information. Applications that don't transmit any personal information or send it to the application's owner, represent an absolute minority [2]. Indeed, personal data of users are essential to the operation of mobile applications to provide custom functionality. For example, a location-based application may use our navigation data to locate us on a map and find the most appropriate route for travel. But the use of these personal data represents a potential threat to privacy necessitating a specific protection.

1.1.3 Requirements about Personal Data Protection

Everyone has the right to privacy. Privacy is *"the ability of an individual to be left alone, out of public view, and in control of information about oneself"* [3]. There are the ability to prevent the privacy intrusion and the ability to control the collection or sharing of personal data. What is meant by data or personal information? *"Any information relating to an identified or identifiable natural person, referred to as "data subject" - an identifiable person is someone who can be identified, directly or indirectly, in particular by reference to an identification number or*

to one or more factors specific to his or her physical, physiological, mental, economic, cultural or social identity" [4]. The processing of these information is prohibited but mainly they may be necessary in some cases, especially in medical diagnostics. But this processing should be clearly defined and must get the explicit consent of the application user. The respect for privacy may be described as *"a right which prevents public authorities or commercial companies from measures which are privacy invasive, unless certain conditions have been met"* [3]. A right to data protection has been introduced and is closely linked to the right to privacy. Data protection principles aim *"to establish conditions under which it is legitimate and lawful to process personal data. Data protection legislation obliges those responsible to respect a set of rules and empowers the people concerned by granting them rights. Finally, it provides for supervision by independent authorities"* [3]. Each country has its own interpretation of these rights, but is based on the he principles have not been defined previously. Today, for mobile applications, personal information may include [4]:

- *"Data collected directly from a user via an application's user interface (name, address, date of birth,...).*
- *Data that is gathered indirectly such as mobile phone number, IMEI*[5] *or UDID*[6]*.*
- *Data gathered about a user's behaviour, such as location data, web-browsing data or the applications used which is linked to a unique profile.*
- *User-generated data such as contact lists, videos and photos, messages, emails, notes, and call logs."*

Mobile applications using this type of personal data need to be protected against all threats to privacy. But today, as shown in the study by Wall Street Journal, to assess the amount of information that is stored on smartphones and which are transmitted by them, *"many apps don't offer even a basic form of consumer protection: written privacy policies"* [5]. This study even shows that of 101 popular smartphone "apps", 56 transmitted the phone's UDID to other companies without users' awareness or consent. This is mainly due to development constraints.

1.1.4 Competition between companies for mobile development

Mobile application development has become over the years one of the largest and quickest revenue generators for companies[7]. It has been steadily increasing over the years. More and more companies have become specialists in this sector, offering their clients a wide range of possibilities in the acquisition of a custom application. But like any commodity, there are between these companies, a constant competition that can be harmful to the consideration of privacy in mobile applications. *"App makers are looking to get their products ready and out the*

[4] According to Article 2 (a) of Regulation (EC) No 45/2001 [3]

[5] IMEI : International Mobile Equipment Identify

[6] UDID : Unique Device IDentifier

[7] The mobile App revenue growth at $25 billion by 2015 (up from approx. $6.8 billion in 2010) according to Market and Markets (Business Market Research Firm)

door so they can derive profit from the highly competitive mobile app market as soon as possible" [6]. The current mobile market is constant evolution in the marketing competition between applications, most designers focuses on the appearance and specific features of applications, which results in the quick development, at the expense of factors such as data protection and privacy. The appearance of an application is essential for the designer because *"the first-open experience is crucial ... This first impression often determines whether the user ever opens the app again"* [7]. For a large portion of developers, an application is designed in this way to persuade the consumer to download it and use it. So, the design will often have a higher priority than data protection.

1.1.5 Risks for consumers and companies

However, giving less priority to privacy may cause some risk, both for consumers or the company developing the application. Navigation data such as *"security credentials, personal financial information, private communications, sensitive company data and more..."* [8] could yet benefit to hackers or malicious people who would use this information to be harmful to companies and consumers. The data loss is the main scourge of companies that develop mobile applications. As says Apple spokesman, Tom Neumayr, *"privacy and trust are vitally important"* [5]. It is also perceived as a risk for consumers of mobile applications with the theft of identity, password, bank account,... that may impact for the user's life.

According to 52% respondents companies to a survey, data leakage is the main concern of businesses [8]. However, *"only 30% of organizations have a security plan that does take into account the changing risks"* [9]. The main consequences for companies related to this lack of efficiency are:

The image degradation - The first concern of businesses is to protect its image. Indeed, if the customer perceives the misuse of such data by someone malicious, he puts a focus on the company. Which means that even if the company is developing a new application on the market that takes in account the privacy, with these older applications were not satisfactory for the users, it is almost sure they will not even bother to download it.

Loss of user confidence - Poor security of users' personal data can lead simply to the loss of customers. It is not enough that a service is operational and attractive to a customer for him to remain loyal to a product. The vendor must constantly renew the confidence of users towards the use of the application. Otherwise, this could lead to the continued loss of customers.

Loss of shareholder confidence - A poor image or a continuing loss of customers in the market for mobile applications could prevent any shareholder to invest in new applications. Why invest when you know that all applications marketed by the company have seen the emergence of fear from users.

[8] `http://viaforensics.com/education/white-papers/appwatchdog-findings-mobile-app-security-iphone-android/`
[9] traduce from French [8].

From a legal standpoint, in case of loss of information, companies would also undergo lawsuits from consumers which may have an impact on the image and the loss of shareholder confidence. To counter risks when developing, companies are based on measures to protect sensitive information especially by security techniques [9]. The logical and physical architecture of an application affects its security. Its core is its collection and presentation of data. These must be handled properly to meet privacy, financial and legal regulations or guidelines. The application and its environment must therefore be secured.

But is that enough to secure data? Many developers think that to take into account the privacy in a computer system, it is enough just to protect personal information by implementing security techniques. This observation is all too present today in the field of mobile applications. But what is the user's place within the system? Must he know personal data used by systems? Of course, as we have seen, legislation forces data controllers of information to develop ways to ensure users' rights. Protection of privacy is not just a security issue but much more.

1.2 Objective

The main goal of this thesis is to provide adequate tools for developers to include privacy early in the design of mobile applications. Principles of privacy being widely studied at present by many organizations, it should not be difficult to propose concepts or methods to be followed in constructing a mobile application when attaching importance to personal data protection. For taking privacy into account in the mobile application development, a specific process must be followed with different properties:

- **Proactive:** to avoid postponing security aspect in the data protection. In contrast to a reactive approach, where changes are performed after a threat or an opportunity has already occurred, the proactive approach promotes the early detection of a potential opportunity or threat. This will allow to rapidly take the necessary measures with regard to respect of privacy in order to avoid any consequences pertaining to risks for a company
- **Agile:** because this type of process is the most suitable solution to the objective, its aim is to supply a product in a short time via the release principle. In addition, Agile processes promise more clarity and flexibility in a project. Since the competition is ubiquitous between companies, the addition of privacy will have virtually no impact on development time.
- **Cross-application:** because the process should be adapted for a part of mobile applications identified by Gartner.
- and mainly tailored to mobile development requiring a specific method.

1.3 Scope

Knowing that it is difficult to propose an unique development methodology for all mobile applications identified by Gartner and to facilitate the research work, we will reduce our study to a limited number of applications. Banking applications such as `Money Transfer`, `Mobile Payment` and `NFC` banking pose no current challenge towards privacy. The banking sector is aware that it should protect personal data of their customers because that could otherwise cause financial loss to the company. But all companies are not necessarily aware of the problem and then forget to protect the users' privacy of their applications. Regarding social online services such as Facebook, Twitter or others, the protection of personal data is unfortunately neglected by designers in comparison to app appearance. The same goes for geolocation systems and targeted advertising [10]. The process that will propose will therefore be mainly targeted at applications such as:

- **Social Networking and Media** including `Mobile Heath Monitoring`, `Mobile Instant Messaging` and `Mobile Music`.
- **Location** for `Location-Based Services`.
- **Mobile Advertising** and `Mobile Search`.

About `Mobile Browsing` applications, a specific process is probably not necessary for the simple reason that a web browser on a computer or mobile is not different in terms of development.

1.4 Structure of the thesis

The thesis will be structured as follows. In chapter 2, we will begin by defining and describing the state of art for having a clear idea of the key concepts to consider for cross-application development in including the principle of privacy in a proactive and agile way. We will relate the current methods, guidelines and tools taking into account the process properties to offer a new methodology that we will define in chapter 3. Finally, in chapter 4, to show whether the process adapts correctly to a given project, we will demonstrate its application on a case of a mobile advertising application[10] designed for Alcaltel-Lucent Belgium. We conclude this work by discussing the difficulties encountered during the approach and an outline of future objectives to be achieved to improve this scientific work (chapter 5).

[10]Personal project developed during the work experience placement of the academic year 2011-2012.

Chapter 2

State of the art

2.1 About this chapter

To meet our objective, initially, we will make the state of knowledge on the subject. Knowing that the process requires a suitable development methodology, different software development process will be described and analysed. It's the same for methods of mobile application development. Privacy must also be taken into account, we will devote a section. During this state of the art, different approaches will be proposed to implement the future methodology, we will evaluate them to show if they meet properties such as being `agile`, suitable for `mobile` and consider `privacy` but also on criteria such as:

User Satisfaction - Determine whether the approach is appropriate for the project that the user wants to design. If it meets desired requirements such as ease of project adaptation, learning time and understanding of all key concepts.

Maturity - Meaning that the approach in use for long enough has removed or reduced most of its initial faults and inherent problems over the years.

Availability of Support and Guidance - To make the practical use of approach easier, one must have enough information to implement it and practical advices must be provided. In other words, approach documentation should be delivered with the product.

User Community - Refers to a group or individual who shares his experience of the approach that can help users unconfirmed. This feedback is useful for the development and maintenance of approach.

Technical Support - In cases where the documentation is not complete enough to use the approach, a technical assistance may be available to find an adequate solution to the matter.

These criteria will put the finishing touches to the evaluation for proposing a methodology combining strengths of approaches that best meet the objective of this thesis. Each approach will bring profits to the future process, we will detail each and propose a procedure to create the new development methodology.

2.2 Software Development Process

When developing a mobile application, a specific development process must be followed. It is important to choose the most appropriate development method. Today, all development method, sequential, incremental and/or iterative, are based on a common model showing the main phases of software development. This model shown in figure 2.1 is called "Systems Development Life Cycle (SDLC)" or software development process. It is used by systems analysts to create or alter information systems. The process consists of five steps common to all software development methodologies:

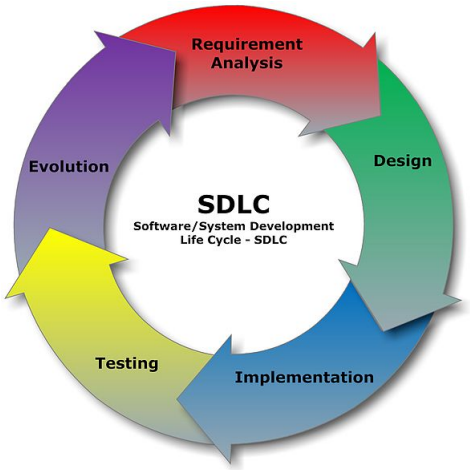

Figure 2.1: Software Development Life Cycle with `Requirement Analysis` as first step.

- **Requirement Analysis:** this first phase is critical to any development cycle, it provides a clear picture of customer expectations about the software to be developed. It includes the definition of project goals through operations and scenarios. Requirement analysis refers from the end-user information.
- **Design:** each goal is described in detail. The designer defines operations with modelling languages, screen layouts for the software interface and business rules. These information are necessary to development team for creating a relevant application.
- **Implementation:** after the design, the development team can write code to implement requirements using the design documents. This step also involves the establishment of environmental systems and definition of test procedures for the next phase.
- **Testing:** when a product is functional, tests must be performed. For this, a testing environment is creating and multiples of numerous tests are performed using the predefined scenarios. If not successful, the software will back to the implementation stage.
- **Evolution:** the last phase of development consists in producing the final product. It is also at this stage that the product is updated.

There are many software development models based on SDLC such as sequential, incremental and/or iterative. The models analysed are best known as the waterfall model (sequential process), spiral model (incremental process) and agile model (incremental and iterative process).

2.2.1 Waterfall model

This model followed by most companies is one of the first formal models of software life cycle. Developed in 1970, quoted in an article by Dr. Winston Royce and based on the life cycle SDLC (described in section 2.2), the waterfall model, shown in figure 2.2 consists of a series of sequential steps where documents are produced at the end of each step to verify compliance before moving to the next step.

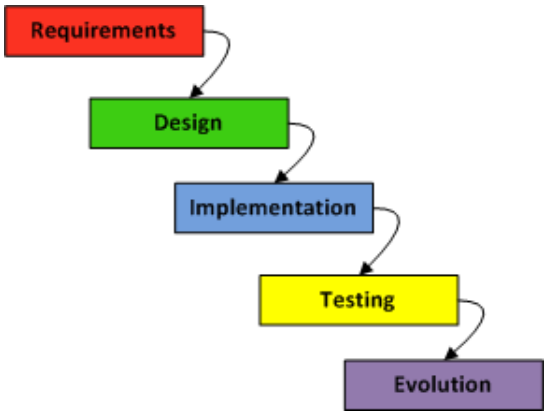

Figure 2.2: Waterfall model

Waterfall discourages revisiting and revising any prior phase once it's complete. This "inflexibility" in a pure Waterfall model has been a source of criticism[1]. *"The criticisms of a non-iterative development approach (such as the waterfall model) are as follows:*

- ***Poor flexibility:*** *the majority of software is written as part of a contract with a client, and clients are notorious for changing their stated requirements. Thus the software project must be adaptable, and spending considerable effort in design and implementation based on the idea that requirements will never change is neither adaptable nor realistic in these cases.*
- *Unless those who specify requirements and those who design the software system in question are highly competent, it is difficult to know exactly what is needed in each phase of the software process before some time is spent in the phase "following" it.*
- *Constant testing from the design, implementation and verification phases is required to validate the phases preceding them. Users of the waterfall model may argue that if designers follow a disciplined process and do not make mistakes that there is no need to constantly validate the preceding phases.*

[1]Identified and summarised in an article on the waterfall model [11]

- *Frequent incremental builds (following the "release early, release often" philosophy) are often needed to build confidence for a software production team and their client.*
- *It is difficult to estimate time and cost for each phase of the development process.*
- *The waterfall model brings no formal means of exercising management control over a project and planning control and risk management are not covered within the model itself.*
- *Only a certain number of team members will be qualified for each phase, which can lead at times to some team members being inactive."*

2.2.2 Spiral model

In order to make the Waterfall model more flexible, Boehm, an American software engineer, developped in 1988 a revolutionary incremental version containing the concepts of waterfall model combined to the features of the prototyping model. Based on an iterative development cycle, with each iteration, the product is increasingly full and solid. The spiral model, represented in figure 2.3, focuses on risk management. It is favored for large, expensive, and complicated projects. For smaller projects, the concept of agile software development is becoming a viable alternative.

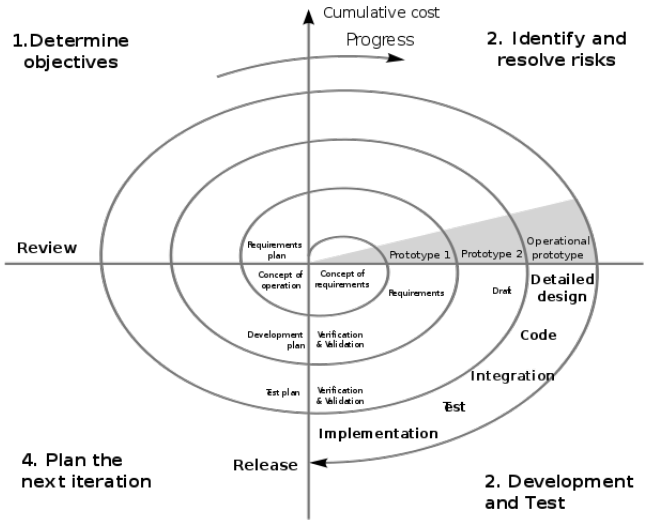

Figure 2.3: Spiral model

Each loop in a spiral represents a development phase. A phase consists of four steps: 1. determining objectives, alternatives and constraints; 2. risk analysis, evaluation of alternatives; 3. development and verification of the solution; 4. results review and plan the next cycle. The advantages and disadvantages of spiral model are as follows[2]:

- *"Estimates (ie. budget, schedule, etc.) get more realistic as work progresses, because important issues are discovered earlier.*

[2]Identified and summarised in an article on the spiral model [12]

- *It is more able to cope with the (nearly inevitable) changes that software development generally entails.*
- *Software engineers (who can get restless with protracted design processes) can get their hands in and start working on a project earlier.*
- *Cost involved in this model is usually high."*

2.2.3 Agile model

Introduced by the Waterfall model designer and based on iterative and incremental development, the Agile method aims shorten the software life cycle, developing a minimum version, by integrating functionalities iteratively via client requirements and all tests during the development cycle. The Agile methods are based on principles defined in a manifest by 17 software developers in February 2001, describing the way of specification changes and modifications will be considered during the development cycle. Through this work, developers have come to value:

- Individuals and interactions over processes and tools.
- Working software over comprehensive documentation.
- Customer collaboration over contract negotiation.
- Responding to change over following a plan.

The manifest does not define in any way an agile method. If the principles are common to all agile methods, how to apply them depends on the nature of the project. The implementation of its principles is done using so-called agile practices. Some of these practices are recognized by experts in software engineering for a long time [13]: 1. Deliver software frequently and regularly; 2. Making short development cycles; 3. Establish a full team for a development; 4. Manage team members by empowering; 5. Have the representative of the users on the same site as the rest of the team; 6. Produce plans at several levels: detailed only for the short term and more general for the medium term; 7. Develop by integrating the code continuously; 8. Make assessments of project to improve the work is done.

Other practices have emerged during the development of agile methods and became indisputable after being accepted by specialists in many projects:

- Have a product backlog and take it into account when defining the priorities (Requirements Management).
- Follow the progress of projects with daily meetings.
- Write tests before writing code (Test Driven Development).
- Practice, sometimes, work in pairs (Pair Programming).

To contribute to product quality and usefulness, agile practices will be combined with manifest principles. There are many official agile methods with adaptations according to the project or company. These methods are not competing with each other, but are specific to the needs. Some are more suited to small teams and other are adapted for substantial team. Actually, Scrum is the most widely used and known to most businesses.

2.2.4 Scrum Methodology

Dedicated to project management, Scrum is mainly used by project managers to develop applications quickly. As stated by Aubry [13], often seen as an agile method, Scrum is not a procedure or method by Ken Schawber, its founder, but rather a framework supplied with a set of elements to perform the project. The framework consists of a team setting up the process, time boxes defining schedules, meetings, journals and retrospective to create consistency, and artefacts to organize the project. The team is composed to a `Product Owner`, a `Scrum Master` and a `Development Team`:

Product Owner - The actor defines the content of the product by collecting the features expected in a list called "Product Backlog" and shares this information with the development team. He is the only one with this power that allows avoiding the dominance of technology in product development and conflicting of interest between participants. The Product Owner must have some skills in order to best ensure his role, he must:

- Have good knowledge of the business domain.
- Master the techniques of product definition (User Stories).
- Be able to make decisions quickly.
- Keep an mind open to change.
- Be a good negotiator.

It is hard to find the ideal person for this work given all the desired skills, but a good product owner will be available, involved in the project, collaborative with the rest of the team and motivated by the task that awaits him.

Scrum Master - This is one of the most important roles of the Scrum process. He is responsible for helping the team to adapt the tool to the context by implementing the concepts defined by the process. The Scrum master must ensure that meetings are properly organized, eliminate barriers to avoid slowing down the team and encourage it to become autonomous. He should:

- Have a good knowledge of Scrum.
- Know the technologies used to better understand the problems.
- Be a good speaker (frequent communication with the team and management).
- Be able to influence and motivate a team.

- Be a good mediator to remove barriers due to conflicts between people.
- Not give up to the first adversity and continue his relentless pursuit.
- Ensure transparency in the progress of the project team.
- Serve the team and be humble.

To choose a good Scrum master, we must avoid a person who manages several teams or who already has a role in the Scrum process (Product Owner) to ensure his full involvement in the project.

Development Team - The team's role is essential as it will look to carry the product through the development of an increment for each sprint. The group is autonomous and has the skills to develop the product effectively. Each member doesn't have a specific role and provides expertise on how to approach a task, the team is multifunctional.

Still according to Aubry [13], the Scrum team makes part of a complete process where many artefacts are involved in the project organization. These artefacts are:

Product Backlog - Depicted as a list, the product backlog contains a set of elements called stories defining clearly and precisely what the software should do. These elements are ranked by priority with respect to the proposed order for implementation by the Product Owner. The backlog is shared between all the people on the team or those wishing to learn more about product development and is not static, it evolves continuously (conducive to change).

Release Plan - It is used to define the sprints to come and their content via the associated stories. The plan is oriented towards customers and users to generate their interest and is updated regularly to reflect changes. It is represented in table format, easy to understand and essential for good communication between project stakeholders.

Release Burndown Chart - To get an overview of the release progress, this chart allows to calculate what remains to be done until the end of the release at the end of each sprint. This will support making decisions on how we continue to achieve the objective and think about the adjustments that could be put in place to ensure this.

Sprint Plan - Defined as a list of tasks, the sprint plan specifies the tasks to be performed throughout the sprint. Each task contains a name and description of work to be done, the story associated with the estimated time for it and the name of the person in charge. The plan will be posted in a visible space for all teams involved in the sprint on a wall chart.

Sprint Burndown Chart - This is a graph, updated daily, describing the progress of a sprint and the tasks it contains. The Burndown Chart has three variants to take in account:

- Burndown of tasks, based on the number of remaining tasks.
- Burndown of stories, based on the number of remaining stories for this sprint.

- Burndown of points, based on the total point of the remaining stories.

"Although it's a good indicator of how much work remains to be done, it doesn't show progress against targets (no preview if a task is finished)" [13].

The Scrum process is complex and hard to define, to understand the concept, it is much easier to explain the mechanics by a simple scheme[3], shown in figure 2.4:

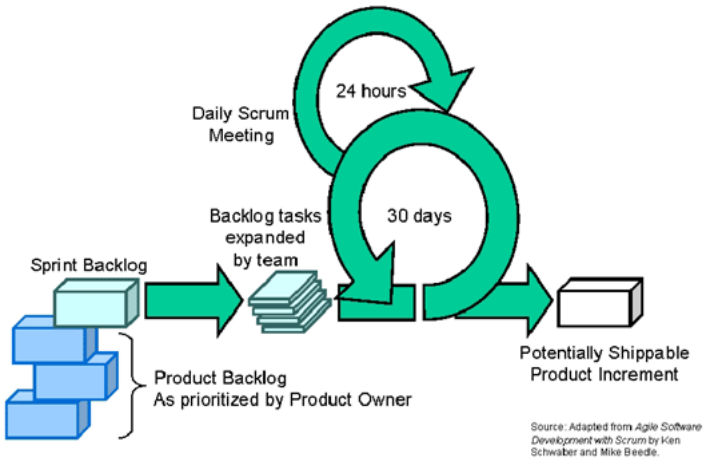

Figure 2.4: Scrum Process

1. The expected requirements are defined in a Product Backlog, they are prioritized and this backlog is managed by a Product Owner.

2. A product version (release or product increment) is developed after a certain number of iterations called sprints. A sprint is a new iteration that produces an increment; its content is defined by the Product Owner and team in order to select new features to implement.

3. Throughout the sprint, checks are made to ensure the success of the sprint. A Scrum Master organizes Daily Scrum Meetings to maximize the likelihood that the team achieves the goals of the sprint.

4. After the end of each sprint, the team gets a partial product, potentially available, corresponding to a release increment. A retrospective (analysis) of the sprint will be conducted to allow for adjustments of the process for the next sprint.

Scrum is an empirical process and made based on the theory of empirical process control [13], represented by three pillars: `transparency`, `inspection` and `adaptation`.

- **Transparency:** all information of the development state must be visible to all those interested in the outcome of the product.

[3]Designed by Craig Murphy for his article "Adaptive Project Management Using Scrum" on the website, `http://www.methodsandtools.com`, consulted on April 2012.

- **Inspection:** in order to detect as soon as possible excessive variations, it is necessary to inspect the development cycle on a regular basis.
- **Adaptation:** adjustments are made to improve the process and avoid future deviations.

2.3 Mobile Application Development

Mobile applications, as any other software, should be developed by following well defined and efficient methodologies. Choosing the methodology for mobile application development is therefore an important decision. How can we do this choice?

Developing a mobile application requires to choose on which hardware and software platforms it will be deployed. Each mobile software manufacturer such as Apple, Google and other had a product line allowing anyone to build a mobile application dedicated to their platform. However, the investment for a mobile application developer can be costly, because he may have to write a separate application for each device. Fortunately, today, several solutions are proposed to improve the portability of the mobile application development. The RDA's business technology consulting[4] describes a number of current techniques of mobile application development [14]. They are divided into three categories: native development, web development or hybrid development.

2.3.1 Native Development

Native Applications use the development tools provided by the Mobile Operating System manufacturer, such as Objective C and iOs for Apple, Java and Android for Google,... Each manufacturer has its specific tool. This type of application must be installed and run locally on the mobile device. Currently, this is the most popular method for mobile applications, for several reasons:

Offline Access - When an application user loses the network connectivity, the native development allows access to local device storage for offline operation.

Device Integration - The native method takes advantage of specific mobile device capabilities such as the camera, GPS[5], Contact List, and network communications. The developer will be totally free to use these services almost without restrictions.

Improved User Experience - The user interface of a native app can be designed to be consistent with what the users expects and is accustomed to on their particular device.

Flexible Synchronization - Native apps can be configured to periodically synchronize with various backend data sources, which can reduce data costs, particularly while roaming. So,

[4] `http://www.rdacorp.com/`
[5] GPS : Global Positioning System

this method provides developers greater flexibility in developing customized database/storage synchronization.

Push Capabilities - Thanks to this capability, we can get important information out to users without having to train them to constantly check the database over the web. Each mobile platform vendor offers a unique push notification service that can only be integrated and employed when developing a solution natively.

Application Market Integration - This is a strategic element in the mobile marketing. It provides distribution of the mobile application via a centralized platform available to all, Android Market for Android, AppStore for iOs,...

Native applications are more present than others on the app stores. They allow using features of the device such as microphone, camera,... for interacting with the components and having a smooth user experience with the smartphone. This is particularly true for applications that take fun pictures, allow talking with another person by video conference, use with the accelerometer for playing fun games,... Native development allows to design applications with better graphics, especially for photo retouching or 3D games. Unfortunately, this method is not perfect and has some disadvantages:

Platform knowledge - The developer must understand the platform operating system and learn new programming languages such as Objective-C for iOs and Java for Android. This can be an obstacle for mobile application development for organizations.

Portability - The major difficulty with native applications is that the transposition of a code developed for one mobile platform to another platform is not easy because certain features behave differently between device platforms. For example, the push notification used by Windows Phone is not the same as that used by iOs and Android.

These disadvantages can be resolved by a new standard. It is now possible with a single code written in one language to develop applications for a range of mobile phones. At present, this method is mainly used for the native applications to reduce development costs and most developers or teams must support apps on multiple platforms. There are more and more tools on the market allowing to make that.

2.3.2 Web Development

Web applications are build with HTML[6] web technologies. They must run into the browser but do not require an installation. This technique is favourable for more reasons:

Less Knowledge - HTML is the predominant markup language for web pages. This language has become normal practice for years and is used by many web developers. Using one technology to create web and mobile application is a considerable advantage. So, the learning is lesser because almost all the web developers also use this language.

[6]HTML : HyperText Markup Language.

Cross-platform - The portability is easier with this method because web applications are accessible from any of various web browsers within different operating mobile system. One web application for several platforms with different operating system.

Low cost development - An unique code allows reducing development and knowledge costs. This advantage is much appreciated by managers wishing to make a mobile application in a short time at low cost.

These applications require less time for the development, because the developer use an unique technology. But they have some drawbacks:

Native Layer - This type of development doesn't allow the developer to access or use the native layer or device specific hardware features. However, an interaction with the device may be necessary.

App Store - Web applications are not accepted in any of the native application stores thereby cutting off an important distribution channel for the application developers.

Even if web applications do not interact with the device, they are more specific for all that is internet content (e.g. online shopping, catalog of products,...). But developers have more difficult to promote them on the mobile application market.

Recently, web applications use the new version of HTML standard. The language groups together HTML5, JavaScript and CSS (version 3 under development) but is called altogether `HTML5`. It brings new functionalities such as new markup elements for new usages, and API to extend the language dynamic possibilities and interact with DOM[7]. The DOM programming interface allows the application to manipulate the content of HTML documents or web pages. It is separated from the platform and language, giving free choice for browser vendors to implement it in their own way. During loading of HTML document, a tree is built via the DOM where each element is considered as an object and can act as a node. This representation allows accessing objects properties and their associated methods.

Despite additional complexity for development and integration, due to the need for downward compatibility[8], the tool is more powerful and this is essential to design an attractive web application. HTML5 definition is still ongoing and is expected to be finalised in 2014 only.

2.3.3 Hybrid Development

To remedy deficiencies from these two methods (see figure 2.5) and create full applications, a method was developed to gather the strengths of both previous methods. This is the integration of mobile web development with native development. Hybrid development enables the use of

[7] `http://www.w3.org/DOM/`

[8] All browsers which recognize HTML5 must too continue to interpreter web pages written with older HTML standards

HTML technology, web development languages, in a mobile application, and it extends native device capability into the mobile web browser.

Figure 2.5: Summary of functionalities proposed for native and web development [16].

Native and Web Development Advantages - With this method, all the advantages of native development such as offline access, device integration, user experience and push notification are present. For web development, the ease of learning and cross-platform development is highly appreciated. An unique standard for all developers allows improving the application longevity.

Hybrid Development Leverages HTML5 - The majority of web applications use the HTML5 technology, this new specification of HTML brings a richer user interaction and capabilities with the set of new functionalities.

Hybrid approach should be especially appealing to enterprise customers who need more longevity in their custom applications. *"It applies the formidable efficiency and agility of web development to the app problem, and is yielding promising results"* [16].

2.3.4 Mobile App Development Guide

Each mobile development method must follow a specific process. There are not a lot of guides for developing mobile applications, one of available is offered by Mobinex company [17]. This approach can be used but not changed without prior permission from the company. The documentation will be used only for editorial and non-marketing[9]. The steps of the methodology are as follows (see figure 2.6):

Needs Assessment - This first phase defines scenarios about how the application will be used in real life. It must take in account a lot of user requirements such as services to integrate, usability scenarios (working environment), targeted platforms, screens resolutions and users who will use the application.

[9]Note that this thesis may be published without requesting permission from Mobinex

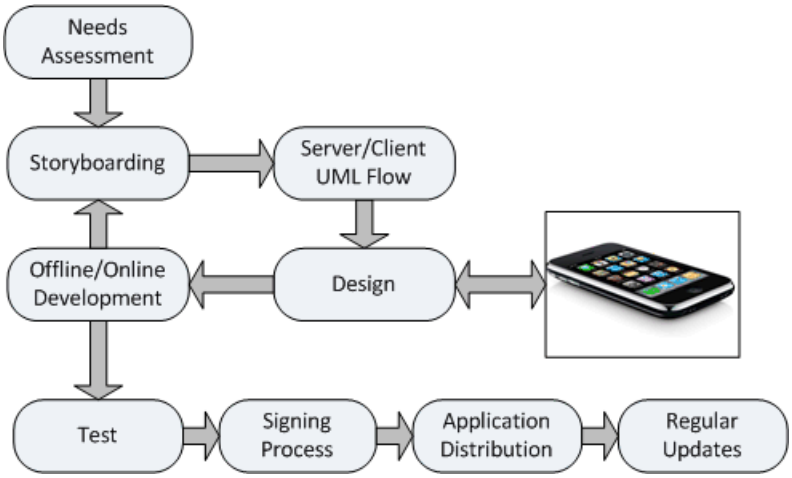

Figure 2.6: Mobile Development Process by Mobinex

Storyboarding - Secondly, the designer must determine the interface and other structural characteristics as the application flow chart, the information included in each pages, the model for the content presentation and what are the types of data (static or dynamic). This phase allows knowing :

- *"The structures of the pages in the application.*
- *Buttons and pages navigation.*
- *Flow of the pages in application scope.*
- *User Experiences on different platforms."*

Server/Client UML Flow - Then, the operations used for data resources and their usage can be described : *"the definition of the resources that will provide the dynamic data, the signature of services that will be used in application (web services, rss feed, json structure) and decision on error code messages (network based error codes)".*

Design - The four phase aims to design the visual appearance of the application interface. The designer will find a name for the application (Brand name and image), determine user experiences for different platforms and types of media used. It's also in this stage that the type of mobile terminal for the application will be defined.

Offline/Online Development - Thereafter, the software environment transforms the offline version to an online version by integrating dynamic data. This step determines the performance criteria of the application. Web services can be needed to deliver data and specific services to the mobile application.

Test - After its development, the application must be tested to find performance problems, bugs, errors,... For that, the developer must write test cases starting at the beginning of the development and will evolve in each step. Additionaly, it is necessary to take into considerations:

- *"Departments who are going to develop the application and test should be different.*
- *Criteria of testing.*
- *What need to be tested with online test tools or with real devices."*

Signing Process - Before the distribution, the application must be signed to run on a specific mobile platform via a specific digital certificate. This step aims to validate if the criteria such as functionality, visuality and/or usability are met by the application. Changes may be made otherwise. This step is necessary that in the case of native applications.

Application Distribution - The next to last stage is the most important because it defines the way to distribute the application to different user segments. The designer can determine :

- *"How will the application be delivered? (WAP Push, SMS Pull, wap download etc.)*
- *How to avoid the difficulties during the application delivery?*
- *How to get reports when distributing application over alternative distribution channels?"*

Regular Updates - An application will ask to periodic updates for improving the services, the design,... One must define what will be updated, by whom and how.

2.3.5 Mobile-D Methodology

Mobile-D[10] is another process to develop a mobile application. It is an agile approach based on XP Extreme Programming, Crystal methodology (focused on the people involvement) and Rational Unified Process (object-oriented). The process is founded upon nine main principles:

1. A project is iteratively where each iteration begins with a `Planning Day` (Phasing and pacing).
2. The `Architecture Line` is a new addition to the already established agile practices.
3. `Tests` are automated.
4. The `Quality control` is applied (Continuous Integration).
5. `Project metrics` are collected to improve processes.
6. `Two programmers` work together at one workstation (Pair Programming).
7. The `Post-Iteration` principle is used rigorously to improve the development process (Agile Software Process Improvement).
8. The customer participates in `Planning` (Off-site Customer).
9. The `identification of end-users` needs must be totally taken into account (User-Centered Focus).

[10]Embedded systems methodology, with as reference book [18].

Mobile-D is composed to five principal phases: `Explore`, `Initialize`, `Productionize`, `Stabilize` and `System Test & Fix` (see figure 2.7).

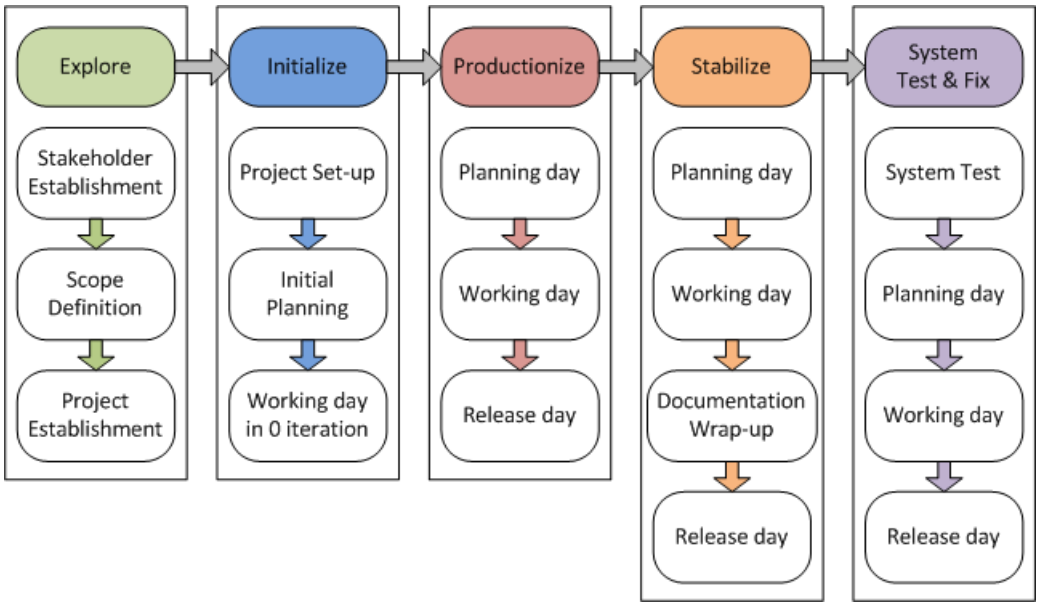

Figure 2.7: Phases of Mobile-D Methodology composed of steps [18]

The `Explore` phases aims to plan and establish project characteristics. Several stakeholder groups are needed to give their expertise about it and define a product line. Three steps are necessary to this phase:

- **Stakeholder Establishment:** establish the stakeholder groups and client.
- **Scope Definition:** determine initial requirements and the timeline of the project.
- **Project Establishment:** plan the project regarding environment, personnel and process issues.

After having completed the previous phase, `Initialize` will aim to prepare and verify all critical development issues to ensure the following steps proceed with respect to customer requirements. This phase comprises three steps:

- **Project Set-Up:** set-up the resources for the project (physical and technical), train the project team (learning languages) and establish a way to communicate with the customer.
- **Initial Planning:** analyze initial requirements, refine and detail architecture line descriptions to create an architecture line plan.
- **Working day in 0 iteration:** implement some main functionalities of system and resolve development critical problems without writing the source code.

The aim of the next phase `Productionize` is to implement the customer prioritized functionality into the product by applying iterative and incremental development cycle. This is done in three steps:

- **Planning day:** plan the work contents for the iteration and generate acceptance tests for the requirements.
- **Working day:** implement the system functionality allocated for the iteration by using the methods of Test-Driven Development, Pair Programming, Continuous Integration and Refactoring. A view of the progress will be provided to the customer to discuss how the software should work.
- **Release day:** Ensure that all of the subsystems are successfully integrated into a single system, that the software is ready for the Acceptance testing and release, and that everything has been done right in the current iteration.

The fourth phase `Stabilize` consists in ensuring the quality of the project implementation and finalize its documentation. This behaves the same way as the previous phase but additionally contains a step of redaction. The aim of the "Documentation wrap-up" stage is to finalize the software architecture, design and UI documents. Ideally, they must be understandable, short, salient, cohere with the source code and be useful.

`System Test & Fix` is the last phase of the methodology, it aims to see if the produced system implements the customer defined functionality correctly, provide information of found defects and fix them. First, the product is tested to detect the defects and document them for the purpose of the fix Iteration. Then, a new iteration is performed to fix bugs but no new functionality is implemented. This iteration behaves like the `Productionize` phase.

When the product has fulfilled all of the phases successfully, the customer will have a first version of it including the priority functionalities. The next feature will be part of another release and so another version will be produced. This principle allows the customer to provide an attractive product for users.

2.4 Privacy

Including privacy in the mobile application development asks to respect principles of data protection. Several people are involved in the process of data protection such as data subject, data controller, privacy commission and processor [3]:

- **Data subject** is *"the person whose personal data are collected, held or processed"*. It is the consumer of mobile applications who has the desire to use a service but is obliged to provide some personal information to use services.

- **Data controller** is *"the person or administrative entity that determines the purposes and means of the processing of personal data on behalf of an institution or body"*. This is representative of the company wishing to develop a mobile application as part of its business.
- **Privacy Commission** is *"the entity which ensures that personal data are handled with care and thoroughly protected, and that future privacy also remains guaranteed.commission"*.
- **Processor** is *"a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller"*. They are people who acts under the authority of "data controller" and is therefore subject to these instructions given directly or indirectly.

The principles of data protection in relation to privacy are based on five points [19][20]:

Transparency - The data collection must be done transparently. Indeed, the data controller is forced to inform the data subject on: information that is collected about him, the one who will have access to their data, how it wishes to obtain this personal data and how the person has the right to rectify their data. It is therefore necessary to provide them with information about the data controller (name, address, ...) for accessing to these rights. The data controller also has an obligation about the data he wish to collect. He has to notify the processing to the Commission of privacy protection.

Identifying Purposes - Before collecting, using, managing and communicating personal data, the data controller must meet a set of conditions and must set goals for the use of these data, such as the storage time before erasure or anonymity. This allows the individual to know the fate of its data and to control them more easily. The data controller should in no case divert from its original purpose and can only process the data collected if it meets at least one of these conditions:

- If the data subject gives his consent to this processing.
- If processing is necessary for the performance of a contract.
- If the law requires disclosure of certain personal data.
- If the processing is of vital importance to the data subject (medical care)
- If the processing should be performed as part of a public interest.
- If the data processing takes an interest for the responsibility or other person, but without counter the interests of data subject.

Necessity/Proportionality - The data controller must limit the data processing to those which is necessary. *"Personal data may only be collected if it is necessary to achieve the previously announced purpose, and if it is relevant."* The data controller must choose the least intrusive path of privacy.

Legitimacy - Data collection and use by the data controller must meet the expectations of the individual and the right to privacy. The data subjects have the right to:

- Be informed about the data processed for knowing the reasons of their processing.
- Ask questions to the data controller to be aware of data he has about them. The data subject must be informed of information the data controller owns and why, what type of data and who will receive it.
- Directly access to their data to be aware at all times of these being collected by the data controller. This right allows them to be informed about the data that the controller collects. This allows knowing the origin of data and the reasons for a given decision. To exercise this right, the data subjects should contact the controller via a letter with a copy of identity card or via mail with an electronic signature.
- Indirectly access to their data. Indeed, some data can not be accessed freely because they were kept under legal protection. In this case, for knowing whether personal data are processed by the controller, the data subjects should contact the Commission by sending a letter with a copy of their ID card. The Commission will then serve as intermediary, consult or modify information in place of claimants. It will inform them after the processing without divulging the data contents.
- Rectify their data if information is inaccurate, incomplete, redundant or prohibited. They can rectify, erase or prohibit these data free of charge. Indeed, the data controller is obliged to consider such requests quickly. Otherwise, a complaint can be sent in writing to the Commission from the data subject.
- Object on the use of their data, but with a good reason to do. The data subject cannot oppose to data processing imposed by a law or a regulatory provision. But for marketing use, he may oppose free of charge and without reason to this treatment.
- Not to be subject to an automatic decision. The decision can not depend on a machine and therefore the law prohibits the processing doing in an automated way. But however, this prohibition does not apply when a contract has been entered or where imposed by a law or a regulatory provision. The data subject also have the right to make his opinion about the data collection.
- Submit a complaint to the Commission or to a court. When they have difficulties to exercise their rights or if they notice that the data controller does not meet his obligations, the data subject may lodge a complaint to the Commission that will try to intervene as a mediator and come to an amicable settlement.

Security/Confidentiality - To meet these obligations, the controllers must implement systems to ensure data quality, allows customer to submit complaints,... This development must be achieved in two ways. On the organizational side where the controller must raise awareness among people who process data relative to the privacy protection. On the technical side, where it will manage access to these data and protect them against erasure, deterioration,...

In general, businesses should consider the fundamental insights that data privacy can offer to organizational security, namely [21]:

1. **Data Privacy is Comprehensive** - it applies not only to the data itself but to the entire environment in which that data is collected and used, principles are provided to help the organizations improve their privacy practices and policies. These principles are called `Fair Information Practice`, described in section 2.4.1.

2. **Data Privacy is Personal** - the interests of the data subject must be considered and built into information systems and controls using a specific approach such as `Privacy by Design` (see section 2.4.2).

3. **Data Privacy Enhances Security** - by minimizing collection, use, disclosure and retention of sensitive personal data, `Privacy Enhancing Technologies` can contribute to stronger data security and allows customizing applications without user identification.

2.4.1 Fair Information Practice

To provide adequate data protection, government agencies of the United States, Canada and Europe have attempted to define information practices by studying the way in which entities collect and use personal information. These good practises allow providing adequate privacy protection. A series of reports and guidelines represent widely-accepted principles concerning fair information practices. These documents are based on five guidelines of privacy protection [22]:

1. **Notice/Awareness:** As advocated by transparency principle, consumers must be informed of the information collection practices of entities. Without that, a consumer cannot make an informed decision as to whether and to what extent to disclose personal information. A notice must be created giving some information about data collection. This document should also identify the rights of consumers to determine whether he has received a right of access to data and how he can exercise these rights in the event of disclosure. In most cases, this information is clearly accessible to the consumer in order to have effective notice what happens to personal information that is inviting to disclose.

2. **Choice/Consent:** the user has the option to consent to the use of personal data and how these will be collected. Traditionally, two types of choice/consent regimes have been considered: opt-in or opt-out. Opt-in regimes require affirmative steps by the consumer to allow the collection and/or use of information; opt-out regimes require affirmative steps to prevent the collection and/or use of such information. The choice can be easily performed by simply checking a box on the screen in a form that will indicate the user's decision to allow or not the dissemination of this information.

3. **Access/Participation:** it refers to an individual's ability both to access data about him or herself.

4. **Integrity/Security:** to ensure data integrity, means must be established to protect against loss and unauthorized access, destruction, use or disclosure of data. The data

will be anonymous, prevent unauthorized access to data, limit access and use only reliable sources of data. Adequate security measures can be taken to prevent access to data such as encryption of communication, data storage on a secure server or use passwords to limit access.

5. **Enforcement/Redress:** the basic principles of privacy protection can not be effective if there is a mechanism in place to enforce them.

2.4.2 Privacy by Design

This concept is *"a significant shift from traditional approaches to protecting privacy, which focus on setting out minimum standards for information management practices and providing remedies for privacy breaches, after-the-fact"*. Privacy by Design (PbD) offers a way to integrate the privacy throughout the development cycle of an application from the project beginning by a proactive approach. PbD was first developed by Ontario's Information and Privacy Commissioner, Dr. Ann Cavoukian, in the 1990s and is based on 7 Foundational Principles [23]:

1. **Proactive not reactive:** *"the Privacy by Design approach anticipates and prevents privacy invasive events before they happen. PbD does not wait for privacy risks to materialize, nor does it offer remedies for resolving privacy infractions once they have occurred. In short, Privacy by Design aims to prevent them from occurring".*

2. **Privacy as the Default Setting:** *"PbD seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. If nobody does anything, their privacy still remains intact. No action is required on the part of the individual to protect their privacy - by default, it's built into the system".*

3. **Privacy Embedded into Design:** *"the concept is embedded into the design and architecture of IT systems and business practices. The result is that privacy becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality".*

4. **Full Functionality - Positive-Sum, not Zero-Sum:** *"Privacy by Design seeks to accommodate all legitimate interests and objectives in a positive-sum "win-win" manner, not through a dated, zero-sum approach, where unnecessary trade-offs are made. Privacy by Design avoids the pretence of false dichotomies, such as privacy vs. security, demonstrating that it is possible to have both".*

5. **End-to-End Security - Full Lifecycle Protection:** *"Privacy by Design, having been embedded into the system prior to the first element of information being collected, extends securely throughout the entire lifecycle of the data involved — strong security measures are essential to privacy, from start to finish. This ensures that all data are securely retained, and then securely destroyed at the end of the process, in a timely fashion. Thus, Privacy by Design ensures cradle to grave, secure lifecycle management of information, end-to-end".*

6. **Visibility and Transparency - Keep it Open:** *"Privacy by Design seeks to assure all stakeholders that whatever the business practice or technology involved, it is in fact, operating according to the stated promises and objectives, subject to independent verification. Its component parts and operations remain visible and transparent, to users and providers alike. Remember, trust but verify".*

7. **Respect for User Privacy - Keep it User-Centric:** *"above all, Privacy by Design requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options. Keep it user-centric".*

Including data protection throughout the development in a proactive way allows reducing the risks for privacy violation. This preventive approach is the heart of many debates where many actors attempt to demonstrate the usefulness of such a concept within information technologies. Recently, a workshop was held in Paris to discuss various issues on the subject. According to the workshop report [24], for organizers: *"Personal data constitute the essential element of the information society, the surveillance society and the digital economy as they are at the heart of both the electronic identification (e-ID), social networks, and devices for monitoring and tracing, whose main characteristic is to be more and more intelligent".* Protecting privacy is essential for strategic actors involved in the current technologies. More and more people get conscious and concerned about privacy issues. Thus, *"the Privacy by Design becomes a techno-legal pilar to guarantee the protection of personal data and privacy of individuals".* Many legal actors having real importance to this principle hope that *"it will be incorporated into EU law, as a principle for any institution or organization, public or private, for which personal data are an important functional and strategic resource".* Here is why, it is important to include this principle in the new development process. In addition, PbD presents some motivations [25]:

- **They can reduce money for development** - *"The cost of including privacy at the system design stage is much less than the cost of having to amend a finished system to make sure it complies with legal requirements and respects individuals' privacy."*

- **They help to reduce risks** - *"Privacy controls that are incorporated into electronic information systems to supplement organisational procedures help to provide additional safeguards which better protect individuals' information from human error."*

- **They help to build trust** - *"The use of privacy enhancing technology in systems helps to signal the integrity and intention of organisations regarding the information that they hold, and encourages trust in those organisations by citizens and customers."*

2.4.3 Privacy Impact Assessment

To manage risks by a Privacy by Design approach, guidelines were defined in a framework called "Privacy Impact Assessment" (PIA) [23][26]. This is *"a process that helps departments and agencies determine whether new technologies, information systems and initiatives or proposed*

programs and policies meet basic privacy requirements. In addition, It also assists government organizations to anticipate the public's reaction to any privacy implications of a proposal and as a result, could prevent costly program, service, or process redesign". The purpose of "Privacy Impact Assessment" is to ensure that privacy principles and legislation are considered throughout the life cycle of a program and even more (see figure 2.8).

Specific goals of a PIA include:

- *"Building trust and confidence with citizens;*
- *Promoting awareness and an understanding of privacy issues;*
- *Ensuring that privacy protection is a key consideration in the initial framing of a project's objectives and activities;*
- *Identifying a clear accountability for privacy issues so that it is incorporated into the role of project managers and sponsors;*
- *Reducing the risks of having to terminate or substantially review a program or service after its implementation in order to comply with privacy requirements;*
- *Providing decision-makers with the information necessary to make informed policy, system design or procurement decisions based on an understanding of the privacy risks and the options available for mitigating those risks; and*
- *Providing basic documentation on the business processes and flow of personal information for common use and review by the department's staff and as the basis for consultations with stakeholders, specifications, information privacy procedures, and communications."*

Figure 2.8: Goals of Privacy Impact Assessment [26]

The framework is similar to a continuous risk management approach. It allows identifying and quantifying risk elements to the project and their potential impact. The process is composed of four steps, `Project Initiation`, `Data Flow Analysis`, `Privacy Analysis`, `Privacy Impact Analysis Report` as described below (figure 2.9) :

Step 1 - Project Initiation - First, the institution must determine whether a PIA is required in projects that raise issues of privacy. For a positive answer, where some programs ask transformations, a PIA is required. Otherwise, it is not necessary to perform this kind of process.

1. **Define scope of PIA process:** the institution must determine whether it has enough information to perform a full PIA. If it just does not have sufficiently detailed information and is not certain that the initiative poses a risk of interference with privacy. The institution must make a *"Preliminary Privacy Impact Assessment"* identifying the types and volumes of personal information, clarifying the roles, responsibilities and policy authorities and determining which aspects of the program

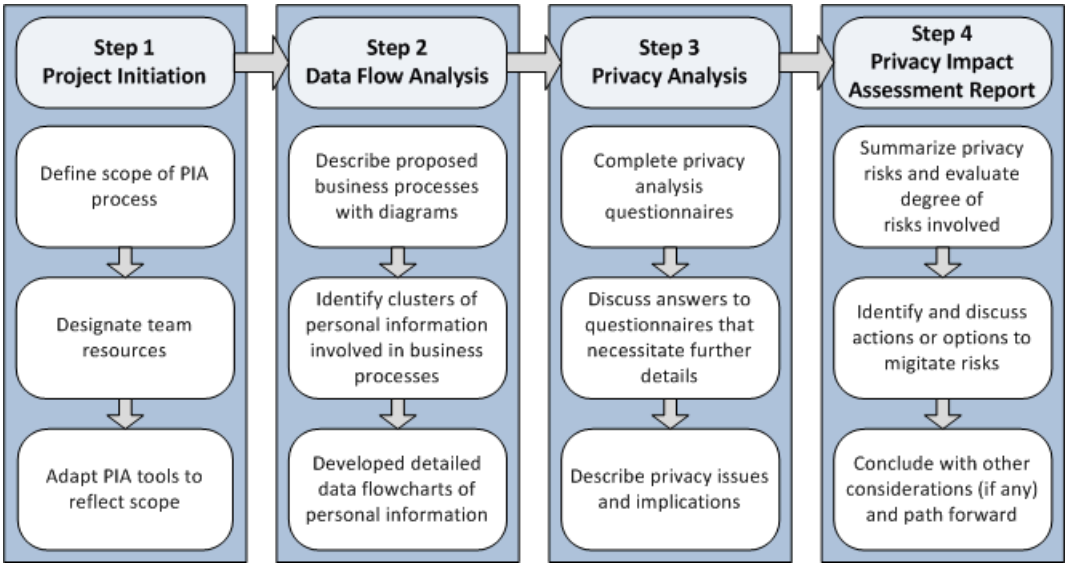

Figure 2.9: Privacy Impact Assessment Process [26]

or service are likely to involve privacy risks.

2. **Designate team resources:** after collecting all necessary information, the institution shall specify the resources necessary for the establishment of this analysis. The team may include legal and privacy experts, as well as program managers or system and IT managers.
3. **Adapt PIA tools to reflect scope:** at the end of this first stage, PIA should be adapted to the project scope. PIAs are not rigid and inflexible tools and two PIA will never be identical.

Step 2 - Data Flow Analysis - After initiating the project, institutions must identify all personal information associated with projects. This activity includes the description and analysis of `business processes` (high level description of the information flows associated with a given activity), `system and infrastructure architectures` (set of security mechanisms that prevent improper access to personal information or maintain any required separation) and `data flow tables` (how and by whom personal information will be collected, used, disclosed and retained) considered for the proposal.

Step 3 - Privacy Analysis - This step analyses the factors related to privacy by reviewing the data flow in terms of policies and applicable laws regarding the protection of privacy. It uses questionnaires as checklists to help identify the obstruction risks to privacy or weaknesses of the proposal.

1. **Complete privacy analysis questionnaires:** institutions must conduct evaluations using questionnaires. There are two types of questionnaires, those applicable to federal programs and services and those applying to government initiatives.
2. **Discuss answers to questionnaires that necessitate further details:** after

completing the questionnaires, additional information may be required to meet PIA objectives.

3. **Describe privacy issues and implications:** this evaluation will be used as a plan to help institutions designated to describe the potential impact at a risk of privacy.

Step 4 - Privacy Impact Analysis Report - The last step is the most critical part of the Impact Assessment to privacy. This is a documented risk assessment about obstruction of privacy and their potential impact. The document includes all methods to mitigate these risks.

1. **Summarize privacy risks and evaluate degree of risks involved:** with the questionnaire, institutions must summarize all possible risks. Notably, they can use a summary table.
2. **Identify and discuss actions or options to mitigate risks:** then, every risk will be associated with mechanisms suggested to eliminate, mitigate or avoid it.
3. **Conclude with other considerations (if any) and path forward:** since it is impossible to eliminate all risks, they should be managed. In addition, institutions must describe an action plan for moving forward.

2.4.4 Privacy Design Guidelines

For helping the developer to incorporate privacy principles in the design of a mobile application, GSMA [4], an association has recently published mobile guidelines. They aim to follow standard practices in the field by offering advice on how the developer will take into account the personal data matter. Their work is coordinated through the *GSMA Mobile Privacy Initiative* gathering representatives from across the mobile ecosystem. This initiative facilitate the development and looking among all practices that best suit a given case. Guidelines are based on privacy principles, defined in section 2.4, and are specific types of applications: `Social Media and Networking`, `Mobile Advertising`, and `Location`.

Transparency, User Choice and Control

According to the principle of transparency, it is advisable to inform the user about the practices used in the collection of personal information. He must always be conscious of the activity that is done with his information. This will aim to establish a system of trust and allow the consumer to make a choice about the use of an application.

[*TUCC1*] - Do not surreptitiously access or collect personal information Information gathering cannot be done without prior acceptance of the user to enable making informed decisions about using an application or service. A form containing facts about personal data manipulation must be provided, this one will describe:

- *"What personal information an application will access, collect and use.*

- *What personal information will be stored (on the device and remotely).*
- *What personal information will be shared, with whom it will be shared and for what purpose.*
- *How long personal information will be kept.*
- *Any terms and conditions of use affecting a user's privacy."*

The consumer can refuse to install an application on his mobile phone if he deduces a dishonest or not acceptable use of his data. This will require to *"ensure usability and avoid excessive user prompts that will burden the user."*

[*TUCC2*] Identify yourself to users - Users should be informed of the source that collects information and how they can exercise their rights or communicate with it. Information such as the name, common name or country of the company must be clearly indicated.

[*TUCC3*] Let users exercise their rights - Provide a genuinely informative privacy statement for specifying how the user can obtain a copy of the information stored about him by the organization for eventually correcting and/or updating them.

[*TUCC4*] Minimise information collected and limit its use - Do not abuse in the collection of information being not need for the application, it must be confined to a restricted part conforming to the expectations of the user. The information collected should not be excessive.

[*TUCC5*] Where necessary, gain the user's active consent - Generally before any installation of a mobile application, the users will give their consent for the access of personal information in several cases:

- If the collection or use of personal information is not necessary for a first purpose, the user must anyway be able to authorize access to its information even if that's for a second use.
- About the sharing to third parties, he will be informed quickly of data that is shared over who will possess them and for what purpose and may at any time refuse the use of such information by third parties. A mobile app should normally provide links to those third parties and their privacy notices to ensure this principle.
- It also happens that the information is stored after use. The user must in all cases be aware of the retention period of these data and how he will exercise his rights in case of dispute.

[*TUCC6*] Give users privacy control - The user should be central to the management of personal information, to become aware of the implications that come up about his privacy, so as to improve his experience for any new perception about the use of new applications.

[*TUCC7*] No silent updates - In case of change that may have an impact on the user's privacy, the data controller must not forget to inform him in order that he takes the necessary steps for continuing or not to use the application thereafter.

Data Retention and Security

In one aspect of personal data retention, the developer must ensure that they are properly protected during storage or transmission. It must think about why the data controller need their information and for how long. Identification is important to focus on measures that will be needed to secure the data from the storage until the deletion. For storage, it is always advisable to avoid keeping personal information too long to avoid risks or costs in case of theft or compromitting of these data. This model is a good compromise for companies wishing not to alter their reputation in the field of mobile application development.

[*DRS1*] Actively manage identifiers - About unique identifiers, it is requested to take measures to ensure that the identifier is associated with one and only one person, application user and is updated regularly. Using identifiers will require to:

- *"Ensure any unique identifiers apply to only one unique user.*
- *Ensure unique identifiers are kept up to date and kept only for as long as necessary to fulfill the applications purpose and reasons notified to users.*
- *Prevent a unique identifier being associated with another user unless required by a justified business need."*

[*DRS2*] Keep data secure - To avoid corruption of personal data in case of unauthorized access, the app designer should ensure their protection when they are sensible such as phone number, mobile identifier, bank details. The technique is to secure these data for the storage and transmission.

[*DRS3*] Authenticate where security calls for it - The user authentication should be considered as soon as possible via appropriate methods.

[*DRS4*] Deletion information - Personal information which is used must be justified, depending on the needs, legal obligations or business model. In case the information is no longer necessary, care should be taken to destroy it or anonymize it, to prevent anyone from finding any trace of a user through these. It is recommended to remove any information that will identify a specific individual.

Education

So that users become aware of how they should manage their privacy and protect their personal information, it is necessary to develop ways to inform him of the dangers they risk in providing such information.

[*E1*] Educate users about the privacy implications - This is to help them by giving them information on the security settings of the application in order to ease the management of their privacy through the available mobile capabilities. Users will need to have details on how they should protect their privacy, preferably in a clear, non technical language

and if necessary be directed to resources (security pages) proposed by more competent like initiatives.

Children and Adolescents

Currently, a lot of children and teenagers use mobile applications. But, they may not be fully aware of the consequences of disclosing their private information and allow others to collect information for their use. Applications specifically used by minors must ensure that *"the collection, access and use of personal information is appropriate in all given circumstances and is compatible with national law and any applicable regulatory codes."*

[*CA1*] Tailor applications to appropriate age ranges - Mobile applications need to be adapted to the user age range and provide means for that minors to understand the consequences of installing an application or service. The risks for the use of personal information from a child or adolescent must be taken into account previously. In addition, language and style of the application must be appropriate to facilitate understanding of installation.

[*CA2*] Set privacy protective default settings - In order to protect minors against all risks of privacy violation, applications must define default settings for the categories of personal information. These settings will prevent users to automatically post their precise location data. For example, this data will be limited to a generic level such as the city or area.

[*CA3*] Comply with laws on the protection of children - Applications specifically used by children and adolescents must comply with rules concerning the collection and use of their personal information as applicable jurisdictions may impose to protect minors. In some cases, parental consent is required before we can collect personal information about minors.

[*CA4*] Verify age where possible and appropriate - When necessary, the age verification may be performed. This access control aims to minimise inappropriate collection and sharing of personal information relating to children and adolescents. In addition, the guideline says that:

> *"Where integration with access controls is not possible, users may be asked to self-certify instead – in that case, they must be asked for a date of birth during installation, activation or registration. Consider that children and adolescents are adept at bypassing safety controls. If users enter a date of birth indicating an age where they must be denied access to a service or otherwise restricted, they must be prevented from starting over and entering a different date of birth during that session and thereafter if technically possible."*

It is also important to ensure *"not to include prompts to the user that could be seen as encouraging them to lie about their date of birth."*

Accountability and Enforcement

An important point is that contributors having an impact on the mobile applications ecosystem must create the tools and interfaces that make these guidelines possible. These contributors include everyone involved in the development, promotion and sale of these applications collecting and using the users' personal information or allowing third parties to have the takeover in these data.

[*AE1*] Assign responsibility for ensuring privacy into the product - Each entity that collect personal information must ensure a company representative is responsible to develop ways for integrating the user's privacy into applications, services and business processes.

[*AE2*] Give users tools to report problems regarding an application - Users should be able to report issues of violation of privacy or security they encounter when using the mobile application. An application must contain information on how they can report such problems. For example, by providing a brief statement and a link to the website of the company clearly and concisely. In addition, *"if the data controller collect email contact addresses (with permission) he could also email that information to users."* To conclude, It is recommend to establish and maintain procedures *"to deal with such reports and address any specific threats and risks."*

Social Networking and Media

On internet, it is now possible to communicate with others through social networks and this context to provide personal information. Because these applications can have a significant impact on privacy, the application developer should take the necessary steps so that the user is informed of data to be manipulated by the application and transmitted to the public. The user should be able to have the choice to engage to this service type, being aware of information that is transmitted and how he may exercise its rights in case of breach of privacy.

[*SNM1*] Prompt users to register for social networks - The user will need to register before using the service to possess a profile on the network. But we must be careful *"not automatically map user registration information to the user's publicly available profile unless the user has been made aware of this and given the option to exercise choice and control."*

[*SNM2*] Ensure default settings are privacy protective - Regarding the data confidentiality, users should be aware of information that will be shared about them in a clear and transparent way. Whether they are shared publicly or not (only visible to the user), they must have the ability to easily control their profile and know:

- *"What information or categories of information will be published about them upon registering.*
- *How they can easily change any default settings.*

- *Whether they and their personal information will be searchable by or alerted to other users.*
- *How they can make private data visible only to authorised parties."*

[*SNM3*] Take measures to protect children - The children's privacy require more protective measures. This can be explained by existing laws or codes regulatory asking to designers to establish specific protections for people with a certain age. In general, *"children must be prevented from publishing contact details or their exact locations."*

[*SNM4*] Create appropriate tools for data deactivation - Methods of deactivation or deletion of accounts or personal data must be set up. The user must be able to remove any personal content stored or handled by the application.

Mobile Advertising

With the increasing number of regulations due to the use of personal data in the advertising, studies have shown that the user must understand how advertising is chosen. The mobile application should give users a way to control this targeted advertising so they feel more comfortable with it. *"Key to successfully driving mobile advertising and ensuring that ads are relevant and useful to users is establishing best practice based on real transparency and meaningful choice and control."*

[*MA1*] Inform users about advertising features - Applications wishing to set up advertising should inform users before they download or install the desired application. The guideline gives advice such as:

> *"The developer could let users know by providing an ad icon and/or a short notice. The icon or notice could link to a URL taking users to more information that helps them understand what information will be used and what choices they have over the advertising."*

[*MA2*] Capture appropriate agreement to target advertising to a user - Users should be aware that their personal information is used for advertising. If this is the case, they must give consent before any manipulation by profiling from applications or third parties. It is recommended to provide instructions on how to modify or delete their profiles and how to opt out of profiling or targeting.

[*MA3*] Target based on legitimately collected data - Advertising is allowed only in cases where personal information is necessary and consistent with the main purpose of the application. In addition, *"targeting advertising to users based on information collected from the handset (for example, location) or the user's interaction with other apps or with the internet must only use information that has been collected in the course of providing the features and functionality the user has requested."*

[*MA4*] Respect privacy when viral marketing - A mobile application is not allowed to send messages to users' contacts without the active consent of the user. It is mandatory to *"respect the privacy of users' network of contacts"*.

[*MA5*] Ensure content is appropriate The advertising content must be adapted to the age of the user. Advertisements must comply with applicable laws, codes or regulations.

Location

The location-based applications are many on the mobile market and generate tremendous interest in the field of privacy. The use of location data is particularly important for mobile users. The choice of sharing the location must be desired by the user and this choice can not be forgotten.

[*L1*] Inform the user that location will be used and give them choice - For an application to locate the user for a location service, it is not necessary to have his active consent (as consent is clear and implicit in the user's request). By opposition, if the user's location is used to build a profile and to target him with advertising, then the user will be notified and shall give his active consent before information is used and collected. So, *"applications must provide clear notice, before a user's location is accessed or collected , about:*

- *What location data an application intends to access (e.g., cell ID, GPS, village or town).*
- *How the data will be used.*
- *Whether data will be kept and how long.*
- *Who location data will be shared with."*

[*L2*] Capture appropriate consents to use location data - If an application wants to retain a history of the user's location, it must notify the user about why and how long the data is held. The user must give active consent, he may view the data and remove them. About applications sponsored and financed by advertising using location data, they will have to set up means so that the user is made aware of this use and need the user's active consent. Generally, for applications that continue to collect, use and share location data during and after using the application:

- *"Users must provide active consent to the continued operation of the location feature.*
- *The application must include a means that alerts the user that the location feature continues to operate via a detected symbol.*
- *Once an application is closed it must not collect location data unless the user has agreed to this."*

The user must have the possibility to switch `off/on` any transmission of location data. This feature should be readily available and offered by the application.

2.5 Evaluation

After reviewing and describing relevant approaches for achieving our objective, we can now evaluate them. Table 2.10 summarises the evaluation of these approaches. For a totally comprehension, we will detail each of them with a few clarifications. Values *Yes* and *No* correspond specifically to the fact that the approach supports sufficiently the corresponding property. The assessment information comes to a personal judgement supported when it is needed by other external bibliographic.

	Privacy	Mobile	Agile	User Satisfaction	Maturity	Guidance	User Community	Technical Support
Scrum Methodology	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Mobile App Development Guide	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Mobile-D Methodology	No	Yes	Yes	Yes	Yes	Yes	No	No
Privacy Impact Assessment	Yes	No	No	Yes	Yes	No	No	No
Privacy Design Guidelines	Yes	Yes	No	Yes	No	Yes	Yes	No

Figure 2.10: Evaluation Table

Scrum Methodology The Agile process does not take into account the privacy. It can be suited to the mobile context because its development cycle is not fixed and can therefore be adapted to any kind of project. Many current projects developed by companies follow this process because it satisfies the expectations of developers. Scrum is well suited for projects requiring a rapid development over time. It became so popular for its maturity, many users have probably experienced problems, but these were resolved long ago. But too because there are a lot of documentation on the subject. The framework is constantly promoted by associations whose mission is to increase its awareness and its understanding. Scrum Alliance[11] is the best known and offers technical support. It also offers its subscribers the chance to be part of the user community. In conclusion, Scrum is well established at present in the mobile software development community and it complies with the requirements.

Mobile App Development Guide This process seems to be a good way to develop a mobile application considering the key concepts of embedded applications development. It includes the same steps as a software development process (SDLC) with additional steps specific to mobile application development. But it does not take privacy into account and is not specifically agile. Indeed, the process must be adapted for cycles of incremental and iterative development or

[11] `http://www.scrumalliance.org`

continuous such as the waterfall model. But, the development guide is not mature enough to have good feedback on the approach. While Mobinex[12] offers a community platform, opinions are not much pronounced on user satisfaction. Still, personally, if we look closer, the process is easily understood with the description paper [17] and should not pose major difficulties in its use. Then, the developer can always use the online support offered by the company in case of problems.

Mobile-D Methodology Specific for the development of mobile applications for iterative projects, Mobile-D was designed long ago for research in a Scandinavian university by an experienced team that defends agile concepts. But, it doest not include privacy and is not quite popular in the industry and remains rather academic. Mobile-D is not used in business anymore, because it was designed for old mobile standards such as J2EE. It could however be adapted to the new mobile standards but the method is static and imposes a process to be followed carefully. An adaptation is therefore complicated and disproportionate. Nevertheless, for the older standard, the approach satisfies the expectations of J2EE developers. The documentation [18] is complete and comprehensive enough to not need the user community and the technical support. This is why these two criteria are not supported by Mobile-D. Although personally, we can see some difficulty in using the process requiring a lot of learning.

Privacy Impact Assessment This maturity approach is an excellent framework to manage risks of privacy in software development. It is not specific to mobile but can be, for the simple reason that the method is particularly interested in requirements. It is therefore appropriate for projects whose risks are sufficiently numerous and well suited to the principle of Privacy By Design. Unfortunately, it is quite complex to implement because the process is too subjective, particularly in the description of business processes where no recommendation is expressed on how the work must be done, but additional documentation can fill in missing information. In addition, development team must pass through a stage that will analyse the current laws on the subject in each country concerned by the development. Workshops are organized for users to give their opinion but not everyone can participate. Community around PIA is not quite present on the net to read about these advices. But this does not change its good reputation, existing since 10 years. [27]. The approach has always been used by people wishing to integrate privacy in their work initiatives.

Privacy Design Guidelines These recommendations help to consider privacy when designing a mobile application. They are specific to mobile but do not refer to an approach determined by a development cycle. It is an aid for those wishing to include privacy in the design of a mobile application. The organization offers a set of rules that can be transposed to a large number of application types. Nevertheless, some rules require more thought and must be adapted to reflect the desired context. The documentation is fairly extensive as shown in the documents present on the organization website. Although guidelines have been provided only recently, the reputation

[12] `http://www.mobinex.biz/`

of GSMA is well established. The association was formed in 1982 and helped develop the world of mobile[13]. Their community includes more than 18500 industry executives. Even if it does not offer technical support for the mobile guidelines, the association can always contacted through social networks.

2.6 Procedure

Considering this evaluation, each approach possesses the key concepts in the implementation of the new methodology. It is therefore easy to combine these different elements to create a process consistent with our expectations. The steps of this new process will be those of the software development life cycle (SDLC) which we will combine progressively new ingredients:

- Scrum can work for mobile application development and is compliant to agile practices. All elements of the framework will be included in the new methodology. It will be the central pillar in terms of team management and development. `Release Planning`, `Sprint Planning`, `Daily Scrum`, `Sprint review and retrospective` allow to supervise sprints during the life cycle of the project.

- About the mobile part, the agile process will be combined to the Mobile App Development Guide proposed by Mobinex. It is complied to new standards of mobile applications, this is not the case of Mobile-D. This second, despite being agile, is a process too old to meet the objective. Adaptation to new standards would be a waste of time. Mobinex offers a ready-made solution where steps are combined as follows:

  - `Needs Assessment` step will be the requirement analysis.
  - `Storyboarding`, `Server/Client UML Flow` and `Design` steps will be located into the design phase.
  - `Offline/Online Development` step and `Test` step are alike to implementation and testing.
  - `Application Distribution` and `Regular Updates` steps form the evolution phase.

  Signing process will not apply to the future methodology, because we advocate to use an hybrid method for the mobile development, not a native development.

- The development method will be the Hybrid approach because it presents a number of advantages (see section 2.3.3). It is a great option for cross-platform requirement and it enables to create fully native app. In addition, there are millions of web developers who already have the base skill set to build mobile apps with HTML standard. HTML5 is the programming language for the new development process.

- For the consideration of privacy in a proactive way, Privacy Impact Assessment process will be part of the design step. It will aim to detect risks to privacy into the mobile application

[13]GSMA History, `http://www.gsma.com/aboutus/history/`, consulted on July 2012

environment. The `Project initiation` step will be implemented in the first phase of development to assess whether PIA is necessary for the project.

- In addition to PIA process, Privacy Design Guidelines will propose ways to take privacy account during the storyboarding phase. Each guideline must be considered according to the type of application. They will then be considered as tasks to implement for taking into account privacy. As for the case of an application asking user consent where the developer will use a check box to ask the authorization to use personal data.

- Others concepts are also used, for the application users and scenarios definition, for the unit testing management and for the privacy policies declaration. These tools will be declared in the next chapter.

Chapter 3

Development Methodology

3.1 About this chapter

To propose an adequate solution to our problem, we will define in this chapter a new development methodology. This methodology will consist of stages of software development life cycle (SDLC) with detailed steps. We will follow a systematic procedure incorporating the key concepts defined at section 2.6 of previous chapter. Each steps of software development life cycle should integrate these concepts to design a development process for mobile application while considering the requirements for the protection of personal data. The objective of this methodology is to facilitate the task of considering the protection of privacy during the development of mobile applications. In addition, the new methodology should conform to agile practices. Scrum will be the central pillar of this new development process. So, the Scrum Master, the Product Owner and the team are important actors in the process. Others actors involved will include a lawyer for the part related to privacy, a designer for the design part of the application and a security expert for the protection of personal data. An overview of the structure of the new methodology is shown in figure 3.1.

3.2 Requirement Analysis

First in any project, we must define the requirements that describe how the product will behave. This step requires constant monitoring and takes an important place in the development cycle, it is inevitable. As a reminder, it captures requirements from users and stakeholders of the future system. This requires a well structured documentation written efficiently. This documentation is commonly called *specifications*. For this new method, we describe how the *specifications* will be prepared, determine what information will be needed in defining needs, and propose a solution to detect whether the protection of personal data is necessary to the future application. This step is decomposed into two parts, one containing the requirements in the form of scenarios and the other detecting sensitive data in these scenarios.

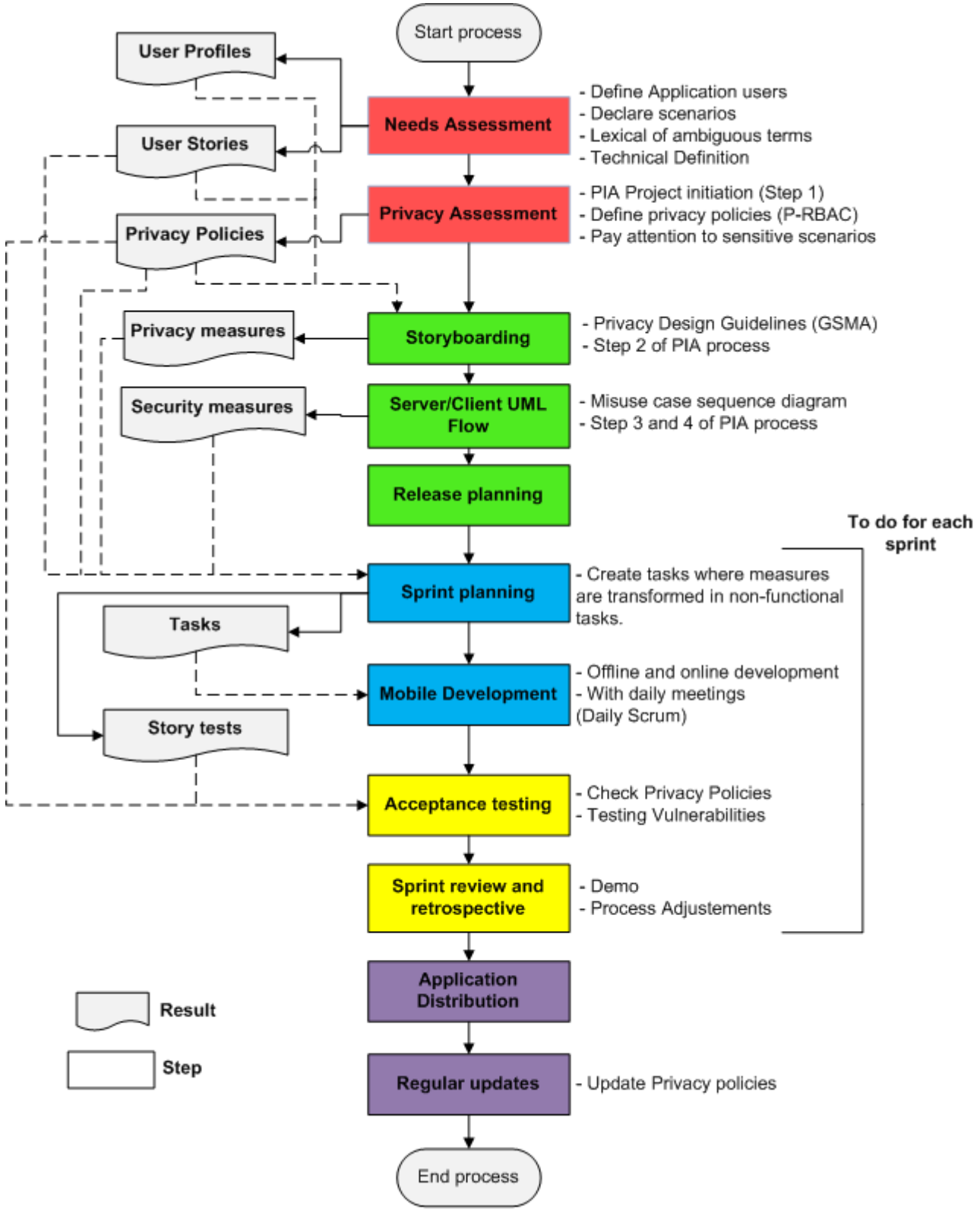

Figure 3.1: Methodology Architecture

3.2.1 Needs Assessment

The requirements capture is conducive to the smooth conduct of the project. These will be described using scenarios. This first part will summarise on the one hand, the future users of the application, scenarios declaration showing how the system should behave and the other hand, technical characteristics of the future application.

Define Application Users At first, for writing scenarios describing system behaviour, it is necessary to identify the users involved in these and categorize them into groups for easier to detect interactions that they may have with the future application during handling information. This step will answer various questions:

- What are the major user groups?
- What are their characteristics?
- What do users want to accomplish using the system?
- What are their overall goals?
- What do users need from the system to accomplish these goals?
- What tasks do users carry out to accomplish their goals?

To get to know the future users of the application, we will observe them, interview them via questionnaires incorporating some of the issues mentioned above. That will detect more easily the goals that the user wants to achieve but also to detect the data that it could handle. Thereafter we can synthesize the needs, tasks and characteristics of different user groups. As proposed by Craig Borysowich, the analysis of characteristics is called Human Factors Analysis [28]. It collects information about users. Information can be of three kinds: demographic (physical aspect), experience (capabilities of user) and others characteristics.

- Demographics include respectively gender, age, physical characteristics such as the fact that it left or right handed or disabilities that may be physical, cognitive, mental, sensory, emotional, developmental or some combination of these.
- Data on experience are the training, prior and current job experience, academic experience and computer literacy.
- Other data may be works habits, preferences (what users like or not), literacy and language skills.

All this data can be useful to better understand the needs of users based on their profile. The specifications should therefore define respectively profiles of each user group. We wish a better view of the sensitive data that the system could handle. These information allows also to select privacy guidelines to consider. Indeed, recommendations are not considered in the same way according to the user age (children or adult). Even physical and mental skills can have an impact on means to inform the user of the risks to privacy. So, we must not discard this categorization.

Declare Scenarios After identifying the users of the future product, based on their characteristic, the needs can finally be defined in writing. The scenarios reporting is an important step for the application development. Today there are several ways to define the project requirements. The method most used is called *use case*, which defines the interactions between an actor (here the user) and the system, to achieve a goal. The actor can also be an external system and must therefore be defined as a user, so it must have its profile as a person handling sensitive data. Use cases are very informative since they determine precisely how parts of the system interact. Yet, in the context of simplification, we will not use this method. Indeed, the principle of *user story* is more appropriate to the agile practices. For efficiency reasons, it is necessary to reduce the step related to requirements, so here is why we will use user stories for defining user needs in a clear and readable way. They are lighter than use cases because they do not require a complete description of scenarios reducing the time spent for specifications.

An *user story* is as its name tells a story in which the user is the main actor. A story should tell something, it must be described on a durable support with a descriptive name and must be standard compliant. To identify the stories, a process of decomposition splits the features of the project in stories. Work sessions should be held in presence of the entire team. Thereafter, user stories will be included in the Product Backlog.

The description of user stories should preferably meet a standard plan. Defined by Mike Cohn [13], recognized as one of contributors to the invention of Scrum, recommends using the following formulation:

As a <type of user>, I want to <some goal> so that <some reason>

- The first part identifies the type of user who wants to achieve the purpose described in the second part of the user story (application users defined in the previous point).
- The second part is the functional purpose to be achieved by the user.
- The last part justifies the goal. However, it is optional, because it is often evident.

As specified by Aubry [13], other attributes may appear in an user story as additional information for development, including UML sequence diagrams but also story tests related to the story. Story tests are black box system tests created from user stories. A story must have at least two story tests: one for success and one for failure. They are written before the sprint beginning (syntax definition in section 3.4.1).

User stories must also go through a stage of decomposition in order to minimize the difficulty of a story that could be complicated to implement. Indeed, user stories must be thin and small (on average carried out in three days) so to be involved more closely in a sprint [13]. There are two ways to decompose an user story. The first is to separate data that are different in nature (eg. bank account or mail account). The second is to separate goals whose actions are different

(eg. create or delete a profile will be two separate stories).

After user stories have been decomposed, they should be classified by priorities for include them in the Product Backlog. The Product Owner and his team will be working together to identify user stories requiring further attention, either because they can be more complicated to develop or because they depend on other user stories with a more higher-priority. During these sessions, the Product Owner should make every effort to convince the team that priorities he proposes are good.

Lexical of ambiguous terms In the definition of user stories, it is sometimes important to agree on ambiguous terms as concepts that can be detected as undefined. For example, for a story in which the action relates to an user profile, it is sometimes difficult to know what lies behind the world "profile". Product Owner may certainly provide full details what is meant by this term. Profile manipulates data such as name, surname, age,... Data that can prove to be sensitive under the protection of personal data. It is crucial to define maximum what is behind ambiguous terms. There are no special methods, a glossary may be used and constantly updated along when new user stories will be created.

Technical Definition After declaring scenarios and identifying users, the customer wishes to develop the system on a set of platforms like iOs by Apple, Google Android,... As we specified in chapter 2, there are three techniques for mobile application development: native, web and hybrid. After comparison, it was apparent that the third method seemed the most appropriate today to develop a business application as quickly as possible. Indeed, it is now easier to design an application without worrying about the native platform containing it. This step is not really important in this new process since it will be specific to hybrid development and the application can run on all platforms available on time today. About resolutions of the future platform, the choice will depend on stakeholder requirements. Maybe that the application should be also compatible for tablets having a more higher resolution.

3.2.2 Privacy Assessment

After defining the scenarios involved in the operation of the new application, privacy concerns can be taken into account by detecting.sensitive scenarios based on "privacy policies" to be defined.

PIA Project Initiation In this first phase, the Product Owner must determine if the future application requires implementation of `PIA` process. Step 1 of this process will apply. The fact that there is data requiring protection allows to say that it is necessary to implement the process. As defined, if there is not enough information to perform a full `PIA`, the Product Owner must make a "Preliminary Privacy Impact Assessment". In addition, user stories will probably be more detailed if they contain personal data. Ambiguous terms should be better defined as precisely as possible to detect and determine whether such data can be crucial to privacy.

Define privacy policies To facilitate the management of personal data, it is important to define the rules for access to such data. Currently, in most organizations, access control models are used in the field of security of information systems. The purpose of these models is to express the rules of permission, obligation or prohibition taking into account the context of the organization. These rules are derived from these three parameters and cover a role in a view for a given activity. The role is the person who acts on the view. The view is the data affected by the given rule. The activity is the action that acts on this data. For data related to privacy, there is now a model called *P-RBAC* defined as an extended role based access control for privacy preserving data mining[1]. This model allows to define rules related to personal data about who can access and what operations can be performed by this person on these data.

A *P-RBAC* rule consists of a role representing the person handling the data, an action (read, write, ...) bearing on these data, the data itself and the purpose of the rule. An expression can be written mathematically to include these elements [29]. However, we recommend to avoid using mathematical expressions for small systems with only a few privacy policies. A formal statement may be sufficient. In case where some problems arise during the drafting of policies, the help of a security expert may be necessary.

Pay attention to sensitive scenarios After defining privacy policies, the Product Owner should indicate whether an user story requires closer attention than another based on the personal data that are manipulated. This procedure aims to facilitate the detection of scenarios that bring difficulties to the development team in the consideration of privacy. The Product Owner will indicate additional information about the level of difficulty for a given story: low, medium or high.

Privacy policies and *user stories* are an integral part of specifications (Product Backlog). We must now consider how these requirements will respect the rules relating to the protection of personal data. The next phase of methodology will be the design.

3.3 Design

The purpose of this phase is to develop the design of the application taking into account the specifications defined above. Most of the time, the mobile application forms a part of a client-server architecture to retrieve data. Therefore, we will identify dialogues between the future application and the server handling data, to detect privacy concerns. It is also in this phase that the designer must find a name for the mobile application and that the full team will create the planning for development.

[1] `http://projects.cerias.purdue.edu/ocrproj/prbac.html`, consulted on August 2012.

3.3.1 Storyboarding

Firstly, to design the look of the future application, we will go through a phase of storyboarding involved in the graphical aspect of the application, not forgetting to take into account Privacy Design Guidelines recommended by GSMA (see section 2.4.4). The designer must indeed meet certain recommendations on the available information on a particular page in the case of protection of privacy. It will base on user characteristics and the type of application to determine guidelines to consider. To help him in this task, he may also ask the legal expert assistance. Then, measures must be defined to implement these guidelines. For this, the designer can rely on user stories that contain the information on the data handled and type of these for taking necessary decisions. But also on the information that flows into the application or outside of it. So, the interactions between system components must be thoroughly analysed. However, it is not mandatory to identify interactions for small stories involving only a few components and personal data. Specially for actions within the application such as the navigation between pages (menu-bar and navigation buttons).

Interactions are represented by flowcharts, in accordance with step 2 of the `PIA` process. Each of these interactions must be analysed to deduce, if for a personal data, there is a risk to privacy. To facilitate this task, we use sequence diagrams showing specifically for each feature, threats that may arise during the dialogue between components. Most of threats occur during the transmission of information between the application (client) and the data server, namely the information circulating outside of the application.

3.3.2 Server/client UML Flow

Considering privacy requires security of personal data. To get to implement this security, designer will first have to describe the flow of interactions between the client and the data server to identify transmitted data and suggest ways to protect them. Each scenario involving personal data must be represented formally through a schema of interactions (flowchart). Sequence diagrams are a good solution but they don't show threats that may prove fatal to data. We will therefore propose a simple solution to represent this information. *V. Katta et al.* [30] have long reflected to the issue and propose an appropriate way of doing that. They have taken the concept of misuse case, used for threat modelling and security requirements elicitation, to invent a new type of sequence diagram where new notations appear, called *MUSD*, for **M**is**U**se **S**equence **D**iagram. In addition to basic notations, attacks are represented through symbols in red representing particular messages and actions coming from the attacker and the vulnerable components of the mobile application (see figure 3.2). This will allow developers to more easily implement appropriate security techniques to best protect personal data in accordance with the security expert's recommendations.

For helping to identify risks or attacks, according to step 3 of the `PIA` process, questionnaires may be also used. These questionnaires should be written using the list of personal data, their

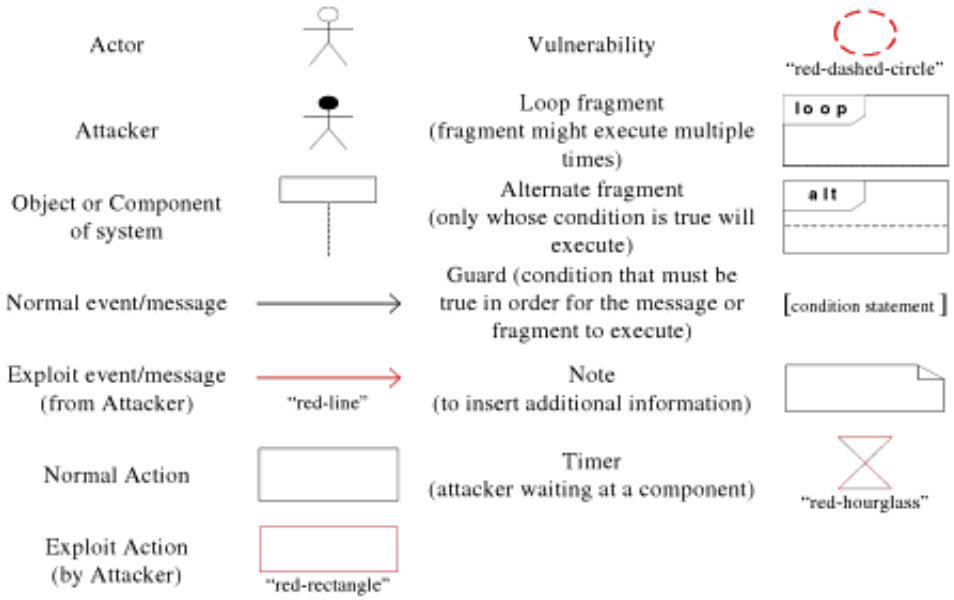

Figure 3.2: MUSD basic notation and its interpretation [30]

associated rules defined by P-RBAC in the previous phase and interactions that exist between the application and the server (represented by sequence diagram). It is likely that the schema of interactions is sufficient to detect potential risks. Questionnaires are not necessarily useful in all cases. Once various risks are identified, the designer will propose measures to try to eliminate, reduce or avoid them. As suggested in step 4 of the `PIA` process, a summary table will do. However, it will pay attention to rules related to the different data involved in the definition of these measures. Each measure must respect the rules associated with data. Proposed measures will be of several kinds and be subsequently applied to development languages. For HTML5 standard, we propose various security techniques during the implementation phase.

Measures to protect privacy will be transformed in non-functional tasks and associated to different user stories to implement. They will not be considered as new stories. In addition, in order to better identify threats for personal data of users involved in these stories, sequence diagrams and measures on privacy and data protection should be part of the project specifications.

3.3.3 Release Planning

After defining the design completely, we can move on to development. But first, it is essential to define the project planning in accordance with the Scrum process. Indeed, this phase is essential; It provides information to those concerned about the content of sprints belonging to the release of the product. Planning for release is done using the Product Backlog defined precisely by the Product Owner and requests the participation of the whole team (Scrum Master, Product Owner and experts included). A meeting is planned at the beginning of the release and preferably before the next sprint. This means that the release plan is continually reviewed and revised after each

sprint. The release planning follows a defined process consisting of five main steps:

- **Determine the end criterion of release:** There are two ways to define the end of a release, base on the backlog; we can say that the release is finished when it will be empty, or else by setting the end date in advance. In both cases, this will give a clear goal of such release to motivate the team. It is obvious that the consideration of privacy could postpone the end of the release because a sprint will take much longer (additional tasks to implement).

- **Estimate the backlog stories:** It will be important to estimate the stories in order to define the effort required to develop them. There is no technique imposed by Scrum, but a common practice today in business is to set up a collective assessment through a `Planning Poker`[2]. Estimation in a group session, with cards, combining expert judgement and estimation by analogy. The effort will be much higher for user stories containing personal information because the measures to be taken regarding privacy are much more important. This parameter is not negligible in the estimation of backlog stories.

- **Set the sprints duration:** There is no accurate method to estimate the duration of the sprints. Despite that agile recommend to have the same duration for all iteration, there are variations about time relative to the project size, how the team works and its motivation. The current recommendation is to have sprints of two or three weeks. But, the duration of sprints will be more important for the simple reason that taking privacy into account is complex. This duration must increase slightly if the team has a good knowledge of security and privacy.

- **Estimate the team capacity:** the ability to work of team while a sprint must be determined with respect to its velocity, its speed to work. But what if the team does not have common experience? The team can simulate the planning meeting at the first sprint (Sprint Planning) to obtain a value for the team's ability. Please note that this measure contains some uncertainty, it should not be taken as part of commitment but rather to allow designing the initial plan of the release. To improve the working speed, we will advocate the pair programming. Indeed, secure and protect a mobile application against the risks demands to be watchful during development. A more experienced person could accompany a novice developer in the area. The allocation of tasks can be done more easily.

- **Finalize the release plan:** After making the entire previous steps, it remains only to plan the release. Given the Product Backlog estimated previously, the release will start with the first sprint by associating the highest priority stories and by adding their size until the sprint limit corresponding to the team's ability is reached. The allocation of sprints will also be based on the difficulty of implementation. For sprints where a number of personal data are handled, it is better to assign them to teams with most experienced in data protection. This procedure will be repeated for future sprints with future stories.

[2] `http://planningpoker.com/`

3.4 Implementation

After taking care of the design of the future application, we can finally move to its implementation. Initially, the first sprint will prepare to choose some of stories to develop. Then, we can proceed to sprint while taking into account privacy. The development will be in hybrid mode with `HTML5` language as we have recommended in the previous chapter (see section 2.6). Each sprint will be accompanied by a phase called `Daily Scrum` to ensure the smooth running of the sprint. Let's start by defining `Sprint Planning`.

3.4.1 Sprint Planning

To prepare the team to work conscientiously and collectively, this phase plans the sprint begins. This gives us a good idea on how implementing user stories, identifying the tasks necessary to achieve the goal of the sprint and their estimation. The plan is organized at a meeting before the start of each sprint while the Team, Experts and Product Owner meet and follow the following planning process:

- **Remind the sprint context:** The Product Owner reminds the coming sprint objectives and the suggested architecture to the team. In stating that the project requires a consideration of privacy.
- **Evaluate the potential scope:** The team decides the scope to be considered in the sprint by selecting a set of elements that will be realized. These elements belong to the Product Backlog. The scope is the sum of the size of selected stories and depends on the team's ability (velocity).
- **Define the sprint goal:** The team agrees on the goal of the next sprint described by a clear and precise sentence. The purpose is mainly to implement features while protecting the user privacy.
- **Identify tasks:** This step aims to allow the team to get organized on how to achieve the scope defined by it, then, each story is dissected to identify the tasks necessary for its implementation. The task types are functional and non-functional. Non-functional tasks correspond to the measures to be applied to take into account the privacy. During tasks analysis, it may be that the team wants clarification of the solution with respect to selected stories if necessary by requesting additional information to the Product Owner. All development activities related to a story must be considered for the development but also performing story tests or learning new technologies. According to Aubry [13], each story test is composed of three elements:
  - *"The state of the software before running the test (also known as test precondition or **context**)*
  - *The **event** that triggers the execution.*

– *The state of the software after the execution (also known as postcondition or **outcome**)"*

These tests are formalized textually as follows:

`Given` context `when` event `so` outcome

They must also be compliant with *Privacy Policies* to avoid during testing to perform unauthorized access to data.

- **Estimate tasks:** After identifying the tasks to be performed, the team will estimate the time needed to achieve them, without spending a lot of time. The tasks will be estimated in hours and should preferably be small enough to avoid spending too much time on a task. The ideal is to have tasks not exceeding two days.
- **Assign tasks:** The tasks are not assigned automatically based on individual skills. The team itself decides how to assign tasks, each member decides what he wants done. We will of course assign the stories with a lot of personal data to consider for pair with high experience in privacy, security but also in web development. Each person has his own experience of the situation and can provide advice on the problem. *"Pairs often find that impossible problems become easy to solve when they work together"* [31].
- **Embark on collectively:** The last step is essential; it aims to solicit the full involvement of each member to fully invest in the sprint in order to protect personal data.

3.4.2 Mobile Development

After setting the current sprint, the team will implement the design of the application. For this, it will use a mobile development tool. There are many tools for developing mobile applications with `HTML5`. No tool is better than another because they have mostly same features. At present, it is difficult to choose a good development tool. To make this choice, the team should agree and choose one from a list of tools in the appendix A, taking time to ask the pros and cons of each of these to suit needs. In all cases, knowledge of `HTML5` is necessary, training may need to be organized before the development of the application. For this scientific work, we will use a few concepts of this language. Applications of mobile advertising, geolocation, social networking and media have in common the storage and transfer of data. The new HTML standard offers several ways to store or convey information effectively. But, they have many security problems as will be explained.

Data storage and security User needs to be always connected to the internet for accessing to services, it may happen that the server is not available. Today, HTML5 enables users to continue interacting with Web applications and documents even when their network connection is unavailable. The goal is to allow the application to run for this period, then synchronizing

information during the connection comeback. The developer must therefore take necessary measures to store modified, created or deleted data and synchronize these data by checking their validity. In the past, for storing data, web developers used cookies. Cookies allow us to login automatically to sites we use frequently, such as Gmail and Facebook . But, cookies *"have been a privacy problem and for a while now web sites interested in quietly tracking users have taken advantage of them"* [32]. HTML5 tried to reduce the problem by proposing a new way to store data, the local storage technology.

Briefly, the local storage keeps user data during the active session without time restriction. It is intended for the data memorisation with a longer life span whose the range is not limited to the active session of the browser window; it is not deleted after the window closing. Data are not transmitted to the server in every request and the developer has a programming interface for accessing to these in reading, modification and suppression, in a practical way. However, the new technology is really not more secure than cookies. If localStorage are not necessarily deleted when deleting history, it will track the user even more than cookies. Different recommendations are available on the web to reduce these security vulnerabilities [33]:

- *"Don't use local storage for session identifiers. Stick with cookies and use the HTTPOnly and Secure flags*[3]*.*
- *If cookies won't work for some reason, then use session storage which will be cleared when the user closes the browser window.*
- *Be cautious with storing sensitive data in local storage. Just like any other client side storage options this data can be viewed and modified by the user."*

Unfortunately, these recommendations are sometimes difficult to achieve, since the HTML5 standard is not yet fully secure and different flaws can be found. We must remain cautious when the use of any technical storage by asking questions about the reliability or otherwise of the technique.

Online Tracking Another issue is that cookies have never ceased to be used for advertising purposes. Called `Third party cookies`, they allow advertising business to track web users by recording individuals' browsing histories for ad targeting. But, the `Online Tracking` presents a privacy concern. To prevent that, web users must switch off tracking via a simple technology. Do Not Track[4] allows web users to express whether or not they consent to having their online activities monitored and collated. *"It signals a user's opt-out preference with an HTTP header, a simple technology that is completely compatible with the existing web."* The header currently accepts three values, `1` if the user does not wish to be tracked, `0` if the user wishes, or `null` (no header sent) if the user has no preference. This technology is available in most browsers such as Firefox, Internet Explorer and Safari and end 2012 for Google Chrome. The web developer

[3] supports data encryption and reduces pervasive attacks.
[4] `http://donottrack.us/`

must disable automatically the online tracking to prevent that personal information are tracked by mobile advertising companies.

Distributed Authentication For mobile applications requiring authentication, we can use the principle of distributed management of identities. From an agreement between service providers called identity federation, it is possible for a user to access multiple services through a single authentication (making single sign-on). This particular circle of trust gives the user control over the confidentiality of his data and allows a better control of privacy. Indeed, little information about the user is sent to the site where he authenticates. The transfer of information is limited to a minimum. The user also has the ability to view and control himself transmission of his information. For web applications, `OAuth`[5] uses this principle. It is *"an open protocol to allow secure authorization in a simple and standard method from web, mobile and desktop applications"*. A Javascript library created for HTML5 allows to implement simply and effectively a single sign-on authentication via this protocol[6]. OAuth is more secured than HTML basic authentication. With Basic Authentication, an user name and a password must be provided to access to the application server, and the mobile app has to store and send this information over the Internet each time you use the server. With OAuth, this is not the case. Instead, the user approve an API to access server, and the application does not store the password. However, nothing says that the API does not take advantage of the situation to handle the user's personal data. It is therefore recommended to the mobile application and application server to ensure that privacy policies are met by implementing a verification system in real time. The user will be notified of any theft of information as soon as that happens.

Geolocation As we have seen various systems can be put in place to protect personal data. For geolocation, there are default means of protecting the user's privacy. In order to respect the user confidentiality, a warning message is showed from the browser asking for permission to have access to geographical data. The visitor can refuse that the web application knows his current position. In addition, the World Wide Web Consortium recommend to provide a mechanism that protect the user's privacy [34]. Developers must:

- *"not send location information to Web sites without the express permission of the user. This permission must be acquired via the user interface.*
- *only request location information when necessary and only use this for the task for which it was provided to them.*
- *clearly and conspicuously disclose the fact that they are collecting location data.*
- *... "*

These measures must be fully respected. The mobile application must in no case force the user to agree to give his location without prior consent.

[5] `http://oauth.net`

[6] A simple example of PIN-based OAuth flow with Twitter and jsOAuth is available at this address: `https://gist.github.com/979955`

HTML5 and security Employing hybrid development may bring some benefits as we have expressed during the state of the art. But using HTML5 can bring a lot of risks. Most new features have many defects as those presented above. But if the developer want to use others, he must be aware of these risks. To meet the guidelines of consent and education, the developer can use Advanced Web Forms proposed by the new standard. It offers a complete range of formats: password, URL, email, search, radio button, checkbox, button,... But there also security problems can occur, especially with the form control outside the contents of the tag "form". Injection of forms is a vulnerability specific to HTML5. It will try to avoid using tags like "form" until this problem has not been resolved. Sebastien Gioria[7] has analysed most of the problems coming from HTML5, but certified that these risks will probably fade over time because the standard is not still applicable at this time [35]. *"The first aim of HTML5 is to create a standard secure, open and simple"*. But he regrets that the opening did so at the expense of safety. In conclusion, one should be cautious in the use of a feature of the new standard. It should also sure it is safe for privacy and can easily address through simple and effective techniques.

3.4.3 Daily scrum

Throughout the sprint, a meeting point between team members is held daily. The Scrum Master is responsible for managing the meeting in accordance with the rules, while reminding to the team that its main purpose is to protect personal data. Scrum process advocates to set the appointment in the workspace where a list of tasks belonging to the sprint is displayed. The meeting takes place preferably in the morning and goes like this:

- **Coming together:** The whole team is present around the table of tasks and members do just that during this time.
- **Answer three questions:** Each participant answers three key questions:
  - What did you do yesterday?
  - What will you do today?
  - Are there any impediments in your way?

  The tasks are moved on the chart. About obstacles, it is possible that a person already has the answer to the problem and in this case, it can give away. Otherwise, an impediments list will be defined and solutions will be considered later through external meetings. It is at this moment that the developer may request the team if he has some difficulties in implementing a story composed of personal data. The team can rely on the advice of legal and security experts if an obstacle is present about privacy. Daily meetings facilitate this dialogue. If a story is too complex to implement, it may also require a more adequate decomposition.
- **Decide about objectives:** After reviewing the sprint progress, the team will make a decision on the adjustment of the sprint goal. Is that the goal will be achieved as was

[7]French security expert playing in the whole internet, `http://blog.gioria.org/`

expected? If the team believes it will not achieve the goals of the sprint, an adaptation may be necessary, either because it will be nothing to show for the sprint review or so because it believes that the stories are too complex to achieve (partly because security and privacy measures are more difficult to implement than expected).

3.5 Testing

After the end of each sprint, the mobile application will be tested before to go to production. Tests will be conducted on its operation but also on its ability to comply with recommendations related to privacy. In addition, a demo will allow stakeholders to give an idea about how the future application works.

3.5.1 Acceptance Testing

Throughout development, the sprint is test driven. In a Scrum process, testing is not an activity that takes place only after development, it occurs the end of each sprint. To verify that the story is finished, story tests (declared in the sprint planning) must run on the latest version of the software. An user story will be therefore complete only if it passes all its tests. Each scenarios must check functionality but also privacy requirements. About privacy, the team must check policies, to be sure that no data is manipulated without their knowledge, and test vulnerabilities of security.

Check Privacy Policies Each P-RBAC rules associated to an user story must be checked. If for some reason or another, a rule is not respected, tests will not be validated. An amendment will be necessary to validate rules. For this operation, there is no particular process. A manual check can be used. But to make our job easier, nothing prevents us to code the rules via unit tests for Javascript and HTML notation. One should also develop a system which during the use of the application checks that privacy policies are followed.

Testing vulnerabilities In addition, vulnerabilities identified during the design phase should be tested. From the schema of interactions (MUSD), it is easy to create test cases to verify that no threats may arise. To do this, there are many testing tools available on the market. `OWASP`[8], a open community working on Web Application Security, provides tools to guide people who want to test vulnerabilities in their web applications. There are training platforms, documents setting out the principal risks of an application and how to fix them,... The most famous is `WebScarab`, this is a proxy with many features useful when performing security audits. In addition to offering the user to view requests exchanged with a web server, it is possible to modify these queries, analyse the session ID,... via an user interface. Other proxies for testing security are available as Paros Proxy[9] which is similar to the tool proposed by `OWASP` in capturing conversations between the browser and the server for analysis.

[8] `https://www.owasp.org`
[9] `http://www.parosproxy.org/`

3.5.2 Sprint review and retrospective

After tests are validated, a demo will be scheduled to show what was accomplished during the sprint. The team will participate in the review along with the Scrum Master, the Product Owner and the experts. Stakeholders may also participate; their presence is highly recommended and encouraged. This review will take place in the last day of the sprint and will be organized this way. Firstly, the team will ensure that equipment necessary for the demo is operational, if the test environment is up and verify that a plan on the conduct of tests has been beforehand devised. The Product Owner will bring to mind the aim of the sprint defined at the meeting start sprint (Sprint Planning). He will therefore focus on the fact that the application protects the privacy of the user. Afterwards, the team performs a demo of the partial product by specifying the deployed stories by indicating measures that have been implemented to protect personal data. Attendees ask questions and also give their opinion about the product.

In addition, the team meets to assess how it worked during the sprint and find a way to improve it in the next sprint. Indeed, it is necessary to make process adjustments in relation to the experience during the previous sprint. The group will share his experience and take the time to think about things that went wrong that went well, to avoid the same mistakes and improve the process about new technological advances. Any team will attend the meeting (Scrum Master, Product Owner and experts included). Experts opinion is indeed highly recommended because they are at the heart of the problem of data protection. They know the difficulties for developers to implement security or privacy measures. Their experience allows to avoid making the same mistakes for the next sprint. The team must also give its opinion because it develops the various modules for data encryption, the layout of consent forms,... Then, the audience will propose solutions that can be implemented for the next sprint and will identify actions for applying the chosen practice in the next sprint through a set of complementary tasks to sprint.

3.6 Evolution

After the mobile application has been developed and tests have been made, the developer will consider publishing it and find a strategy so that the mobile user wants to use it regularly. We would define a good way to supply a mobile application to consumers.

3.6.1 Application distribution

The designer must publish the application in an online store with a set of mobile applications categorized according to their purpose. The consumer can make his choice in the store based on four types of information: price, description, screen captures or a video, and user ratings. The publisher can influence the price but not the other elements. User's interest must be aroused so that he wishes to download the application. But, how to arouse this interest?

1. Produce a demo video to host on `YouTube` and embed in social media (blogs, Facebook, Twitter etc) that replicates the user experience. This video must focus on measures that

will protect the privacy of the user. Indeed, the user can perceive that he will constantly involved in the mobile application and he may enforce his rights to data protection. Potential customers and reviewers will make a quick decision on whether to purchase or review the mobile app based on the video.

2. Segment the product into a full *paid* version and a lite *free* version, where the user downloads a base application that is free but with little functionality, or a pay application but with the entire functionality. The designer can make money from both models, the free version gives users a chance to sample the application and hopefully decide to upgrade to the paid version with all features enabled. Free applications can generate revenue through embedded, third-party advertisements by embedding a simple ad bar into the product. For the paid version, it will set a price corresponding to the application functionality. But does it supply a free or paid application?

    - A *free* app retains a wider audience and more users, reinforces the brand image but many downloads are required to make beneficial. But the ad makes part of the system. One might therefore ask whether it does not interfere with the primary goal that is to design a mobile app which considers privacy. Indeed, some advertising uses these applications to collect personal information. We therefore advocate in the case of a free version to ensure that the third-party advertisements comply with privacy policies.
    - Revenues of *paid* app are immediate, there is less risk for privacy and no dependence on advertising but fewer users use the application and requires a strong commitment to quality.

So many marketing methods aim to tempt consumers to download the application. Today, app stores have become the one-way street for developers. Over 45% of the respondents to the Developer Economics 2011 Report carried out by Vision Mobile [36], uses an app store as primary route to the market, climbing nearly 30% since previous year. This observation is shown in figure 3.3. However, the use of app stores as a primary distribution platform varies greatly by platform. The use of app stores is much more pronounced for platforms that have a native app store. The designer does not just put his application in the store for what it can be used by many users. He must anticipate the features that will be appreciated by the consumer in order to be different from other applications.

3.6.2 Regular updates

The designer must therefore continue to adapt its product (new platforms) and develop it to enrich the user experience (provide new features). He does not forget his goal: provide mobile applications that will ensure constant interest from the user.

"Understanding your app's users, their relationship with your app, and the app lifecycle should influence the design of your app." (UX-Matters) [7]

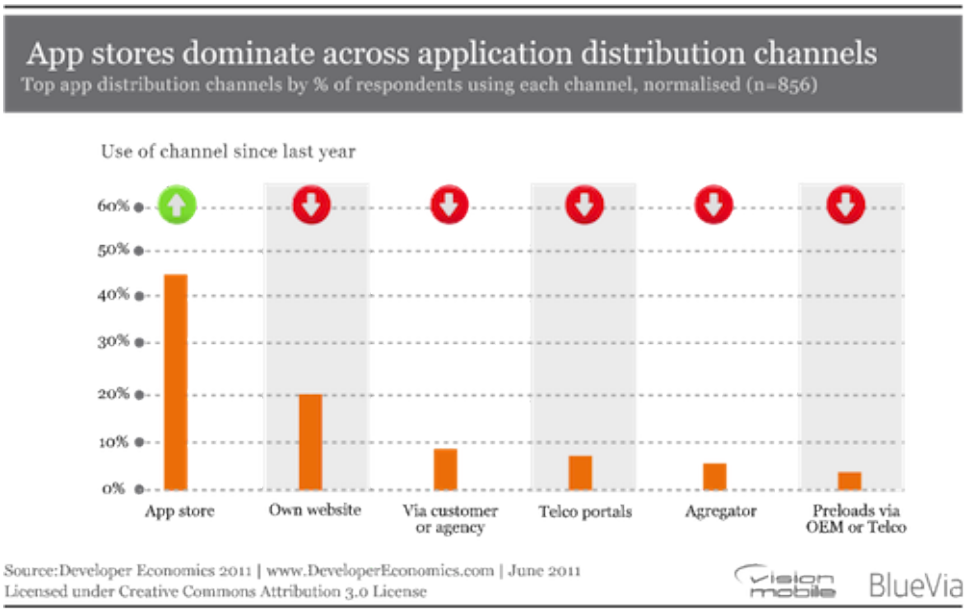

Figure 3.3: App stores have become the one-way of street for developers [36]

When time comes to make an update, it is important to look at the feedback received from users. If enough users are requesting a feature, it is a good idea to look into what would be required to add that feature. However, the most frequent types of application updates are bug fixes. Bug fixes are typically covered in *revision* or *bug fix* releases. Bug fixes don't change the structure or feature set of an app. Instead, these updates make sure that the app is working as designed. Updating an application will be part of a particular procedure approaching the development step. However, about personal data protection, it is perhaps legitimate to update privacy policies. New Requirements surely will represent new stories where new data can be manipulated. It is therefore necessary to go through a complete phase of development to take into account these new updates. These new user stories will simply consider new sprints to develop.

Chapter 4

Case Study

4.1 About this chapter

For testing the development methodology, we will apply it in a particular case derived from a real case, namely the case Alcatel-Lucent (ALU). The concerned application is a marketing software which we developed for Alcatel Belgium and requires a consideration of privacy because it handles personal data. The first part of this chapter will describe the mobile application and what were the expectations of the company for the development. We briefly detail the problems encountered and improvements needed. Then, we describe the application of the methodology to the case and discuss the validity of the approach on the case study.

4.2 A mobile application for Opt-in to advertising campains at Alcatel-Lucent

Two years ago, Alcatel-Lucent has invested heavily in the development of a Mobile Advertising solution (Optism). The solution comprises an Ad Selection Server (ADSS) allowing mobile operators (MNO) and advertisers to create targeted advertising campaigns for mobile phones. To access the system and agree to receive ads, end users must opt in to advertising. They can therefore participate in campaigns. ALU wanted to further enhance its solution with high focus on "usability and end user control". This is why, the project, started in September 2011, was to develop a Web portal where end users could manage and update their opt-in and profile status. When a user opts in to advertising, he will need to specify his preferences. His preferences (profile) are stored in the ADSS and are used to target advertising campaigns. The portal could also allow the user to "view" past campaigns, voucher or rewards he received[1]. Having recently acquired OpenPlug, a mobile software and applications development tools vendor, ALU was interested in a prototype development of a multi-phone application operating as a Web portal. To understand how this application was developed, a description of its functional architecture will be made. At the time of development of the application, privacy was not considered fully in the application development. Therefore, some improvements in this area are possible.

[1] Users will typically always receive a reward from their operator in case they agree to receive ads.

Architecture

With his mobile phone, the end-user can have access to and manage his profile and status. Access to these information is done via the mobile operator that must have an application server accessing to data from the ADSS. Indeed, each mobile operator will have its own application server being able to offer the mobile application to its customers. The smartphone will send request to this server, located in the MNO environment, by the internet network. Dialogue with ADSS and other services offered by the Optism solution will be secured to prevent someone to acquire secret information. The project architecture, represented in figure 4.1, has been divided in three parts:

1. The mobile application developed with OpenPlug.
2. The application server (web service for the mobile app)
3. The Optism solution containing ADSS and other services (also present)

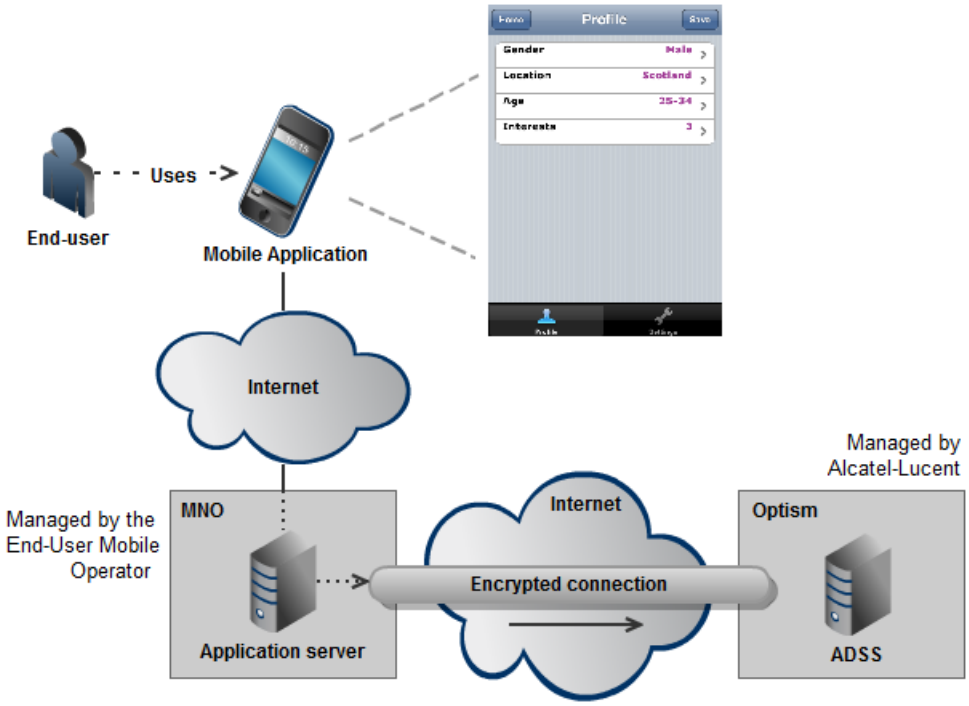

Figure 4.1: Project Architecture

Developed together with the mobile application, the application server receives the set of requests from the mobile and must perform to necessary operations to satisfy customer needs. It works like a web service supplying specific functionalities such as getting and setting member info, making an opt-in or an opt-out,... Each service requires the end-user identification on the application server. So, end-users must be subscribed and authenticated from their mobile application. They are identified via their phone number (MSISDN). Then, user data can be retrieved based on this number from the ADSS. For confidentially reasons, the MSISDN number

is anonymized with a solution called ASI for anonymous subscriber identification. ASI number anonymizes the mobile number before calling services from the Optism platform. This function is available on a web server in the MNO environment. The mobile operator is the only one to know the exact identity of a subscriber. As expected, this function has been introduced to prevent ALU to be liable for any damage caused to the privacy of an end-user.

Improvements

In the current prototype implementation, intuitively, we detect several problems about personal data protection. Firstly, there is no security in the communication between the mobile application and the application server. Transiting data can be subject to external attacks. User privacy is also not taken into account as advocated by legislation. The mobile number travels from and to application server without protection and the end-user authentication is not secured. This is explained by the fact that the application is a prototype. It will be used for demonstration and should not be likely to be available to an outside network. For this scientific work, we imagine that this application should be made available on the mobile market. Mobile operators need to take care that their users do not perceive advertising as spam. End users must stay in control, to protect their personal information and privacy. So, we apply our methodology to develop a mobile application containing the same functionality as the old one but taking privacy into account. The development tool will not be OpenPlug because it is not compatible with HTML5 and it has some functional bugs. But also because Alcatel-Lucent has decided to stop future development with OpenPlug. The tool did not meet expectations of the company for cross-application development.

4.3 Applying the methodology

4.3.1 Requirement Analysis

Firstly, we must identify user groups who will use the application and detect their needs, tasks and characteristics via Human Factor Analysis. By analysing the requirements, one actor will handle information via the mobile phone. It is the end-user who has to manipulate customer data in order to consult or make changes. The user wishes to view his preferences, update them, change his status (opt-in or opt-out), participate in campaigns but also retrieve his historical record (all the messages sent and received in the past). To perform these operations, he should also authenticate on the application using a login and password initialized during a registration phase. The end-user will have different characteristics to perform tasks such as:

- Being able to use a mobile phone (knowing how it works).
- Having more than 18 years (political requirement).
- Understanding English (application developed for this language).
- Having all physical ability to handle the application with his fingers.

- Not being influenced as a person with psychomotor retardation and developmental disabilities (to avoid unwanted consent).
- Being able to judge the stakes of the handling of personal data.

After having identified actors, let's declare scenarios with user stories. For this scientific work, we develop the scenario about the update of user preferences. This scenario has a lot of personal data to handle. It is for this reason that it is not worthwhile to study other scenarios which requires less work in the consideration of privacy. The user story to develop will be the following:

As an `end-user`, I want to `update my preferences (profile)` so that `my personal information is incorrect or non-existent`.

The non-existent term means that the profile is empty probably because it comes to opt out to advertising. And after this operation, the profile of the user is emptied and reset by default. Afterwards, as recommended by the methodology, user stories must be decomposed to reduce the complexity in development. Here, the user story will be separated in two parts because for updating the profile, the end-user must firstly view it for then edit it. This is expressed in this way:

First story (S1)

As an `end-user`, I want to `see my preferences (profile)` so that `I know what it contains`.

Second story (S2)

As an `end-user`, I want to `modify my preferences (profile)` so that `my personal information is incorrect or non-existent`.

In the next step, these user stories must be classified by priorities. Nothing complicated here, the first story (S1) has priority because to update its preferences, it is necessary to know changes to be made. They are by no means more complicated one than the other but the second (S2) depends on the first. The Product Owner will have no difficulty to convince the team that his judgement is correct. He should also provide some details on user stories as ambiguous terms. For these stories, to detect personal data, hiding information behind the preferences term must be described by a lexicon. The user profile contains these attributes:

- **Gender** - refers to the sex as male or female. But, it can also be considered as the gender role in society because it *"describes the characteristics that a society or culture delineates as masculine or feminine"* [37].
- **Location** - refers to the position. It will not be a specific location, the user may choose a country from a list.
- **Age** - refers to the age grade in which the user is. The first grade is 18-25 years.

- **Interests** - refers to things the user loves. His hobbies or preferences such as sport, beauty, culture,... selected from a given list.
- **Language** - refers to the dialect spoken by the user that he will select from a determined list.
- **Time zone** - refers to the region on earth where the user is located. Most of 40 time zones will be suggested according to the UTC[2] standard.

To these information is associated a number which identified the user on the application server. In a database available on this server, this number is linked to end-user mobile number (MSISDN). The identification number, not visible in the application, allows to retrieve information from the application server by sending it as a parameter to services offered by this one. But this number is visible on the network because it is not encrypted. This is also the case of login and password. We see therefore why the user privacy is in danger actually and why we need to apply this methodology. So, the fact to detail the ambiguous terms allows to see a little clearer about data that may cause a risk for the user. To counter these risks, the future application requires to implement the Privacy Impact Assessment process.

The Product Owner has enough information to perform a full `PIA`. As we can see, detailing ambiguous terms prevents any further study. A Preliminary Privacy Impact Assessment is so not necessary because we have enough information about personal data handled. After having collected information, team resources must be specified as recommend in step 1 of the PIA process. The team should comprise legal and security experts for helping designers and developers. Their advice are useful and necessary to facilitate consideration of privacy.

For more accurately assessing risk, as recommended by the methodology, we also need to define who can handle the relevant information because they have permission. Different roles will access preferences. The end-user who uses the mobile application to modify these personal information. But also the mobile operator that can also use its web customer care in case the end-user will be unable to change himself his profile. In the Optism solution, the customer care is the former which allows users to edit their profile. So, we detect two privacy policies for accessing to user preferences:

1. The `end-user` has `write and read access` to `only his preferences` for `updating` ; *changes will be notified and stored in the database of the application server and automatically recorded in the user historic.*
2. The `mobile operator` has `write and read access` to `users preferences` for `updating`; *changes will be automatically recorded in the user historic.*

Following the methodology, we must provide special attention to user stories handling personal data. Here, the first story will not have a higher attention than the second vis-à-vis privacy

[2]UTC : Universal Time Coordinated

because they share the same policies and uses the same information. The end-user and the mobile operator can access in the same way to the information but the second user must realize that he handles personal information of several people. However, the first user story will be higher priority but it can only be explained because the second depends on it. Privacy policies and user stories will be placed in the Product Backlog allowing to design the application. By default, its appearance is suitable for mobile devices with `320x280` resolution.

4.3.2 Design

After defining the product requirements, the designer must devise its graphical aspect, taking into account Privacy Design Guidelines. Compared to the application developed with OpenPlug, there will be no major differences in the appearance of objects. The profile will be displayed in the same way but adapted to HTML5 standard using forms. The management of pages will be similar to the old project but the application will notify the user during the navigation through notifications for privacy. To meet requirements related to the protection of personal data, the designer must propose ways to implement GSMA recommendations. First, he must identify guidelines to consider in developing a Mobile advertising application. They are those related to:

- Transparency, User Choice and Control *[TUCC]*
- Data Retention and Security *[DRS]*
- Education *[E]*
- Accountability and Enforcement *[AE]*
- Mobile Advertising (type of application) *[MA]*

Section about children and teenagers should not be taken into consideration because the application is aimed at adults (over 18 years). Afterwards, he will advise the team on how to implement recommendations. With the help of the legal expert, the designer will draw models for how to integrate privacy notifications in the mobile app and when they should be visible. These advice are measures to implement included in user stories. For example according to the principle of transparency, it is stipulated that the user must give consent on the use of personal data before installing the product. To meet this requirement, a form must contain the facts about the data manipulated by the application and must be shown to the user *[TUCC1]*. When changing his profile, the user must be aware of the data that will be manipulated: age, location,... But also that information is stored on the device (for displaying) and sent to the `ADSS` via the application server after an update. The mobile operator will be able to change preferences only with his agreement. No other person may access them, even another end-user. The form must also specify that the profile is used for marketing only if the user opts in to advertising via the system. Of course, this is a part of recommendations to take into account. While the end-user browses through the application, various notices are announced to explain how he can protect from risks of privacy and how he may exercise his rights whenever necessary. This step is long and rigorous, it requires to put oneself in user's shoes for realizing we can not use his personal

information at our interest without his prior consent.

According to what is proposed by the methodology in the previous chapter, to detect these risks related to privacy, the designer can rely on sequences diagram representing interactions between components of the application. Because it is appropriate to take security measures to protect personal data *[DRS]*, these schema allow to assess where system vulnerabilities are located. In this case, we can use the sequence diagrams developed during the requirements analysis of the previous project. Dialogues between components will be identical where we add `MUSD` notations.

Example of a *MisUse Sequence Diagram* for the authentication operation (see figure 4.2):

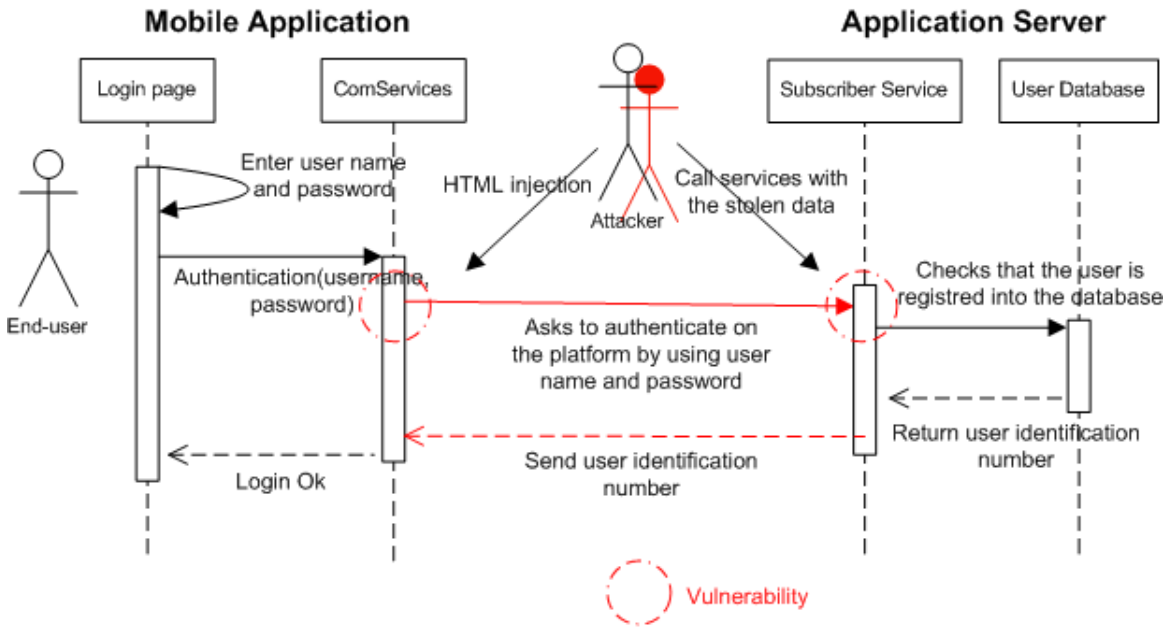

Figure 4.2: MUSD for authentication operation

As we can see, `MUSD` notations show clearly where the vulnerabilities appear. The attacker is located on the internet network between the mobile application and the application server. He has a set of tools allowing to intercept communications and can easily retrieve the user name and his password to make use of the various services provided by the application server. The attacker can endanger privacy of the victim by stealing information such as personal preferences. It is therefore important to encrypt external communications as we explained in the definition of the case study. On the client side, different vulnerabilities are also present. The attacker can pretend to be the application server and inject fraudulent code into the mobile application. However, all vulnerabilities can not be represented by a sequence diagram, it is for this reason that the PIA process completes the risk detection.

To facilitate vulnerability detection, we will put ourself in the security expert's shoes who perfectly knows the risks that may exist from web applications. The expert may base himself on the OWASP Mobile Security Project[3] identifying the top 10 of mobile risks for 2011 [38]. With this list, he can identify threats for privacy and proposes ways to reduce, eliminate or avoid them as recommended in step 3 and 4 of `PIA` process. The only difference is that we do not use questionnaires which are mainly used to help non-experts and may have the impact of increasing the development time which is inconsistent with the agility. Indeed, the time devoted to create and fill them can vary depending on the greatness of application architecture. Some risks identified by the project may occur, we will explain how to mitigate them for the mobile app and describe advice that the expert could give knowing OWASP recommendations:

- **Insecure Data Storage** - To ease matters, personal data are sometimes stored locally on the device requiring special protection. These data are often poorly protected, and can have significant impacts on privacy and information loss. The expert will advise to store only the relevant information for minimizing threats. For an authentication system, it should avoid to record the password in clear as is commonly done. About mobile application, we will encrypt login and password in the case of a remember function exists. The identification number used for the dialogue with the application server will be also secured.

- **Weak Server Side Control** - In order to fully secure the system, back-end services must also be protected. The application server developer must therefore take steps to ensure that its platform and its services are always secured. The expert will propose to equip the server with a firewall, if it is not yet the case, to prevent anyone to scan the network and find a flaw.

- **Insufficient Transport Layer Protection** - The transport layer between the mobile and the server must also be encrypted. Data transmitted over networks are poorly protected and can facilitate Man-in-the-middle attacks. The mobile developer must ensure that sensitive data leaving the device are encrypted using standard techniques such as the user name and password. We pay also particular attention to the personal phone number that travels between the phone and the application server in the case of a registration on the platform. Preferences do not require any security since normally these do not identify one single person, two users can have the same profile. However, we will encrypt them to prevent from being manipulated by third parties for marketing purposes.

- **Client Side Injection** - Mobile applications usually share common libraries. These are sometimes used by malicious person to conduct SQL and HTML injection in order to find security vulnerabilities and gain access to personal information. It is strongly recommended to use prepared statement for the call to database and ensure that dynamically generated page content does not contain undesired HTML tags. When using data retrieved from third-party applications, it should also check the validation of data coming from and

[3] `https://www.owasp.org/index.php/OWASP_Mobile_Security_Project`

going to these applications. So, we avoid the use of unofficial libraries which we have no idea how they were developed.

- **Poor Authorization and Authentication session** - To prevent unauthorized access to the application server, user authentication/authorization and session management must be implemented correctly. Avoid use of potentially compromised values such as IMEI or UUID[4] to identify end-user. As an identifier, it is recommended to use a token session and ensure that tokens can be revoked quickly in the event of lost/stolen device. Solutions exist today when developing web services to allow secure connections. WS-Security is a communication protocol that offers all tools to setup a token session.

- **Side Channel Data Leakage** - There is not a place where personal data are not located. Web caches, screenshots, logs, keystroke logging are only a part. Not disabling platform features and remaining programmatic flaws have a considerable impact on privacy and data theft. The expert recommends a number of recommendations to prevent any threat as never log sensitive data to system logs, remove sensitive data before screenshots are taken, disable keystroke logging and utilize anti-caching.

- **Sensitive Information Disclosure** - One should also differentiate stored and hard coded data. Developers have the annoying habit to hard code important values in their code as API keys, passwords or sensitive business logic. But these data can be easily discovered. Through reverse engineering techniques, we can now retrieve the application code and therefore have access to these information. We advise developers not to hard code data in the binary application.

To facilitate the user authentication and to avoid problems related to the theft of the password or phone number of this one, the development team could also turn to a solution of distributed authentication (described in section 3.4.2) for the mobile application. The end-user would no longer need to register via the application for accessing to his information. Simply imagining that the mobile application asks to the user to send a free SMS to a service offered by the mobile operator that activates his account on the platform. This service would require in return his email account to associate it to his private mobile phone number. The end-user must so login on the platform with his personal identifiers to access his preferences. This practice allows to prevent the sending of mobile number via internet network so not be intercepted by an attacker. A session token would also allow the mobile application to communicate with concerned services. But even in taking all necessary precautions to avoid these risks, other vulnerabilities may occur. For the simple reason that the user must also protect his mobile phone against external attacks. Just another untrusted application installed on the phone can become an entry point for malicious people. This one can so access information stored on the phone. The user must therefore take some personal action to avoid any outside threat (identified in appendix B).

[4]UUID : Universally Unique IDentifier

After defining goals to protect user privacy, the Product Owner must plan the project with the participation of the entire team (experts included). The effort requested to develop user stories is more important than the old project. By considering the measures to consider, we must estimate the development time to 4 weeks. Three weeks for the development of features and one week for the implementation of measures having assumed that the development team has extensive experience in security, privacy and web development. To facilitate our study, we evaluate one sprint including stories that will be described in the requirements analysis. Of course, before the sprint, other stories must be implemented such as allowing the end-user to authenticate to the platform via the distributed authentication system.

4.3.3 Implementation

Before starting to implement the sprint, the team must define a planning. It is to identify and assign tasks necessary to achieve the sprint objective which is to enable the user to update his preferences. Tasks identified by the development team for the first story are:

- Create a web service to retrieve the user profile;
- Create a web page to show and store preferences;
- Call the web service for reading;
- Create a warning message to inform the user that the profile was recovered;
- Design a navigation button to access the preferences; ...

In addition to the functional tasks, we find tasks to be implemented to consider the privacy in a proactive way. These correspond to experts recommendations defined during the design phase:

- Secure web service against any outside attack (WS-service and firewall);
- Design an encryption module to encrypt data transmitted during the call for service (SSL);
- Create a module to protect information stored on the web page;
- Create a notification containing the personal data handled by the application *[TUCC1]*;
- Design a genuinely informative privacy statement *[TUCC3]*;
- Add a link to recommendations on how the end-user can manage and protect his personal information *[E1]*;
- Create a page containing information on how the user can report a privacy violation *[AE2]*;
- Design an icon “privacy violation” that refers to this page;
- Create a message to warn the user that he can opt out to advertising via a button *[MA2]*; ...

This is also at the sprint beginning that tests are declared. Each user story should have at least two tests: one for success and one for failure - Example, with the story "SEE MY PREFERENCES":

Success Test

`Given` the user "John" connected to the mobile application and the application server available and functional.
`When` John click on "view profile".
`So` John's preferences are retrieved and displayed on a page named "Profile".

Failure Test

`Given` the user "John" connected to the mobile application and the application server unavailable.
`When` John click on "view profile".
`So` a message is displayed on the John's mobile phone indicating that the server is unavailable and it is impossible to display his preferences.

It is evident that story could have more than two tests, especially for privacy notices that must appear at the right time. These tests can be functional as non-functional. After identifying the tasks, the team will estimate the time it takes to implement them. It should not spend more than one day to implement a task because the project is oriented pair programming. This solution allows indeed to significantly reduce the development time spent on the implementation of privacy and security measures.

4.3.4 Testing and Evolution

Last steps of the methodology, after the sprint, tests will be conducted on the product version. Each story is controlled without forgetting privacy policies that will be verified. For our case study, we must therefore ensure that the end user is the only one to access information and nobody else besides the mobile operator can access. This rule can be verified easily by trying to access information by unlawful means. The tester can perform different attacks on the server but also on the mobile application where the user is identified. If all these tests are passed successfully, the product can be produced.

About production, `ALU` will incorporate this product like demo version into its Optism solution. Mobile operators may propose this solution to their users or redesigned their own product. It is the latter which will insert the product into the official download platforms, free or paid according to their needs but avoiding any intrusion into consumer privacy by the third parties. The user can then download it and use at will. During usage, the operator must establish means for detecting previously existing outside threats. In case of inadequate protection of privacy by

the application, ALU will proceed to update its product by applying the new methodology. But this could occur rarely because the team took into account experts recommendations during development. Updates will be in most cases of functional type.

4.4 Discussion

By carefully following the new methodology, we came across some obstacles that had to be solved gradually with the development. This section explains what went well and the encountered problems. In the first step, requirements analysis was performed in a simple way because the new application has the same functionality as the old one. The human factor analysis requested a slight reflection to identify the actors involved in the system. However, the procedure was a bit ambiguous lacking detailed documentation on how to proceed. So we created a simple list of the user characteristics. About user stories, the notations allow to easily describe the different scenarios to be implemented. Decomposition and prioritization was not complicated since operations are depending on the nature of the treatment, which in our case depends only on the reading and writing of preferences. After that, the definition of ambiguous terms has not presented a major difficulty. The glossary defines each term precisely. Afterwards, privacy policies have been formally described without real difficulty due to the low complexity of the project to develop.

In the second step, we encountered minor difficulties. Starting with the privacy design guidelines, where it is quite complicated to convert these guidelines into measures without the help of a legal expert. We played the game however to get into the skin of an expert for defining our own measures according to the study case. About risk management, we initially used the old sequence diagrams where faults detected before the application of the methodology were represented. Attempts have also been to apply Privacy Impact Assessment (PIA) as advocated the methodology to detect risks and propose appropriate measures. But after some trials, we came to the conclusion that PIA could significantly increase development time and with the limited information available it is difficult to create one's own questionnaires. It is for this reason that we deliberately deviated from the methodology by relying on a security expert who would base on misuse sequence diagrams and a predetermined list of mobile risks for proposing measures. Knowing these risks, we also detected new recommendations that the user must take to protect his mobile phone from any attack. Finally, in order to plan the project, we have chosen to give an extra week to consider privacy in the sprint. This choice may be the same for any project where the development team already has a good experience in privacy and security protection.

In the third step, the issue was first to define the sprint planning. It took place without any real difficulty. Each user stories can be easily transformed into tasks. The same applies to the measures which were clearly identified during the design phase. The story tests are also easy to create according to user stories and privacy policies. About testing and evolution, it is not easy to draw a conclusion because the mobile application has not been implemented in practice.

Chapter 5

Conclusion

Consideration of privacy into the design of software has now become a necessity, especially in the area of mobile computing. Today, too much personal data is used without user consent and few efforts are used by the developers of these applications to protect data from threats. An attempt was made in this scientific work to propose a new methodology for developing mobile apps by including the foundations of the protection of personal data. Based on the software development life cycle (SDLC), we aimed to integrate the privacy in a proactive way by being compliant with agile practices. To do this, we relied on existing approaches by taking the crucial elements responding to our main goal. Throughout the work, we focused on the integration of these elements. Scrum has been used in particular to meet the agile property, proving to be an incontrovertible solution for project and development team management. For mobile aspect, we realized an inventory of the different approaches that exist on the market today. After analysis, we focused on a hybrid solution combining the advantages of web and native applications. About proactivity, we cannot omit to consider the Privacy by Design (PbD) based on the Fair Information Practices. PbD also provides a way to guarantee that privacy principles are considered during the life cycle of a program (privacy impact assessment). Other initiatives also exist in the mobile world to take into account privacy. This is the case of GSMA that offers a number of rules that can be easily transposed to a large number of applications.

Having identified these concepts, we have evaluated whether to approve their use in future methodology. A procedure has been described in order to explain the reader on how these elements were assembled by following SDLC steps. Various recommendations were defined step by step to help the development team to consider privacy in a proactive way. Then, to validate the process, we applied it to a study case where a number of limitations and difficulties of implementation have been identified. Firstly, it is difficult for a non-experienced team to use the new methodology. Indeed, one can not omit a good training of the team in security and privacy, but also in the project development managed by Scrum. Project leaders must constantly ensure that the team has the necessary faculties to follow such a methodology. The help of an expert in the team can, as we have seen, be necessary simply for a matter of convenience in understanding the various recommendations. Not relying on an expert is also possible but this could have an

impact on the time spent for understanding these recommendations. We also found that the privacy impact assessment can significantly increase development time. This is explained by a detailed analysis for risk detection. The procedure is so complex to follow for lambda users. We tried to address this problem in the study case by providing a simple solution but requiring the presence of a security expert. Based on misuse sequence diagrams (MUSD) and a list of mobile risks, the expert can recommend various measures to prevent any threat for personal data. The methodology has tried somehow to meet the different properties that it should satisfy. We note nevertheless that including privacy in an agile process is not an easy task for obvious reasons. We understand now why mobile application developers have some difficulties to make the step towards a PbD process. However, we hope that this methodology will allow them to realize that it is not impossible by respecting a lot of criteria.

Bibliography

[1] Jerry Kang and Dana Cuff. Pervasive Computing : Embedding the public sphere. Electronic Document, May 2005.

[2] CRIOC. Vie privée mobile: un flux d'informations personnelles difficile à contrôler. Website, September 2011. `http://www.oivo-crioc.org/FR/rssdoc/6189`, consulted on January 2012.

[3] European Data Protection Supervisor. Data Protection. Website. `http://www.edps.europa.eu/EDPSWEB/edps/cache/off/lang/en/EDPS/Dataprotection`, consult on April 2012.

[4] GSMA, Mobile and Privacy. Privacy Design Guidelines for Mobile Application Development. Electronic Document, February 2012.

[5] Scott Thurm and Yukari Iwatani Kane. Your Apss Are Watching You. Wall Street Journal Website, December 2010. Consulted on January 2012.

[6] Steve Tarlow. Study: Major mobile apps compromise your personal data. Website, June 2011. `http://www.newsytype.com/7395-mobile-app-security-holes/`, consulted on February 2012.

[7] Michael Griffith. The Lifecycle of a Mobile App, a User's Perspective. Website, October 2011. Consulted on February 2012.

[8] Bertrand Lemaire. Inquiétude des entreprises sur la sécurité du cloud et web 2.0. Website, November 2010. `http://www.intrapole.com/spip.php?article951`, consulted on February 2012.

[9] NTObjectives Mike Schema. Web Application Exposure to Risk: Raising Awareness to Build Confidence and Improve Security. Electronic Document.

[10] appWatchdog. Improving mobile app security for consumers. Website. `https://viaforensics.com/appwatchdog/`, consulted on July 2012.

[11] Select Business Solutions. What is the Waterfall Model? Website. `http://www.selectbs.com/analysis-and-design/what-is-the-waterfall-model`, consulted on April 2012.

[12] Select Business Solutions. What is the Spiral Model? Website. http://www.selectbs.com/analysis-and-design/what-is-the-spiral-model, consulted on April 2012.

[13] Claude Aubry. *SCRUM, le guide pratique de la méthode agile la plus populaire.* DUNOD, Paris, 2010.

[14] Chris McGuirk, Tony Pekala, Jason Petrin, and Eka Renardi (RDA Development). Choosing the Right Mobile Development Method. Electronic Document, 2011.

[15] Rodolphe Rimelé. *HTML5, une référence pour le développeur web.* Eyrolles, Paris, 2011.

[16] Jeff Whatcott. HTML5 and the Rise of Hybrid Apps. Website, November 2011. http://blog.brightcove.com/en/2011/11/html5-and-rise-hybrid-apps, consulted on April 2012.

[17] Mobinex, leading provider of mobile applications and on-device solutions. Mobile Application Development Methodology V3. Electronic Document, May 2010.

[18] Abrahamsson & al. Agile Development of Embedded Systems: Mobile-D, version 1.0. Electronic Document, April 2005.

[19] Commission for the protection of privacy. Website. http://www.privacycommission.be/en/, consulted on April 2012.

[20] Jean Marc Van Gyseghem. Introduction à la protection des données à caractère personnel. CRIDS (Centre de recherche information, droit et société), 2012. Consulted on April 2012.

[21] Ann Cavoukian, Ph.D. Identity Theft Revisited: Security is Not Enough. Electronic Document, September 2005.

[22] Federal Trade Commission. Fair Information Practice Principles. Website, June 2007. http://www.ftc.gov/reports/privacy3/fairinfo.shtm, consulted on January 2012.

[23] Pat Jeleson and Anita Fineberg. A Foundational Framework for a Privacy by Design, Privacy Impact Assessment. Electronic Document, November 2011.

[24] Francesca Musiani. Compte-rendu de l'atelier anr « Privacy by Design (PbD) . Mettre la technologie au service de la vie privée : Enjeux, limites et perspectives ». Website, March 2012. http://adam.hypotheses.org/1230, consulted on April 2012.

[25] Information Commissier's Office. Data Protection Technical Guidance Note: Privacy Enhancing Technologies (PETs). Electronic Document, April 2006.

[26] Treasury Board of Canada Secretariat. Privacy Impact Assessment Guidelines: A framework to Manage Privacy Risks, August 2002.

[27] Treasury Board of Canada Secretariat. What's the recent history about privacy? Website. http://www.tbs-sct.gc.ca/pgol-pged/piatp-pfefvp/course1/mod1/mod1-3-eng.asp, consulted on July 2012.

[28] Graig Borysowich. User Characteristics - Human Factor Analysis. Website. http://it.toolbox.com/blogs/enterprise-solutions/user-characteristics-human-factors-analysis-18765, consulted on July 2012.

[29] Jean-Noël Colin. Sécurité et fiabilité des systèmes informatiques : matières approfondies - Autorisation et controle d'accès. Electronic document, Avril 2012.

[30] Vikash Katta, Peter Karpati, Andreas L. Opdahl, Christian Raspotnig , and Guttorm Sindre. Comparing Two Techniques for Intrusion Visualization. Electronic Document, November 2010.

[31] Cockburn Alistair and Williams Laurie. The Costs and Benefits of Pair Programming. Electronic document, 2000.

[32] Jim Rapoza. HTML5's Privacy Problem. Website, October 2010. http://www.informationweek.com/news/security/229200574, consulted on July 2012.

[33] Michael Coates. HTML5, Local Storage, and XSS. Website, July 2010. http://michael-coates.blogspot.be/2010/07/html5-local-storage-and-xss.html, consulted on August 2012.

[34] Andrei Popescu (W3C). Geolocation API Specification. Website, May 2012. http://Dev.w3.org, consulted on July 2012.

[35] Emmanuelle Lamandé. HTML5 : quid de la sécurité ? Website, February 2012. http://www.globalsecuritymag.fr/HTML5-quid-de-la-securite,20120215,28468.html, consulted on July 2012.

[36] Matos Kapetanakis. Developer Economics 2011 – Why app stores are a one-way street. Website, June 2011. http://www.visionmobile.com/blog/2011/06/developer-economics-2011-why-app-stores-are-a-one-way-street/, consulted on May 2012.

[37] Ann-Maree Nobelius. What is the difference between sex and gender? Website, June 2004. http://www.med.monash.edu.au/gendermed/sexandgender.html, consulted on July 2012.

[38] Jack Mannino, Mike Zusman and Zach Lanier. OWASP Top 10 Mobile Risks. Slide Document, September 2011. http://fr.slideshare.net/JackMannino/owasp-top-10-mobile-risks, consulted on August 2012.

[39] Jeyaganesh. 6 Best Cross platform mobile development tools. Website, August 2011. http://devlup.com/mobile/cross-platform-mobile-development-tools/2416/, consulted on August 2012.

[40] JPaul Ruggiero and Jon Foote. Cyber Threats to Mobile Phones. Electronic document, 2011.

Appendices

Appendix A

Best Hybrid Development Tools

For helping companies to choose a good development tool, we propose an exhaustive list of tools available on the market. *"These tools are ranked based on the popularity among developers, functionality and ease of use for beginners"* [39]:

PhoneGap is a framework that helps to develop apps for iPhone, iTouch, iPad, Android, Palm, Symbian and BlackBerry devices using web development languages such as JavaScript and HTML. It also allows for access to hardware features including GPS/location data, accelerometer, camera, sound and more. The company offers a cross-platform simulator (an Adobe AIR app), as well as online training sessions for helping to access native APIs and build functioning mobile apps on the PhoneGap platform. More information on `http://www.phonegap.com/`.

Appcelerator's Titanium Development Platform allows for the development of native mobile, tablet and desktop applications through typical web dev languages such as JavaScript, PHP, Python, Ruby and HTML. Titanium also gives its users access to more than 300 social and other APIs and location information. More information on `http://www.appcelerator.com/`.

AppMobi provides development tools and services that complete HTML5 as a mobile platform with game acceleration technology. AppMobi's post-deployment cloud services include user authentication, in-app payments, rich media push messaging, user analytics and live app updates. Is is available in the Apple App Store since March 2010. This framework is even open source. More information on `http://www.appmobi.com/`.

RhoMobile is a mobile development Motorola's solution. RhoMobile Suite platform built to meet requirements of the next generation of business mobility. The developer is free from OS design constraints, able to create business applications that are every bit as elegant looking and intuitive as their consumer counterparts. The solution supports iPhone, Android, Windows Mobile, Research in Motion (BlackBerry), and Windows Phone 7. The product can be free downloaded. More information on `http://www.rhomobile.com/`.

Sencha proposes several products to design HTML5 mobile application. The company was the first to provide tools for developing in hybrid. An app builder allows to deploy apps from a single, integrated environment. A specific framework builds fast and impressive apps that work on iOS, Android, BlackBerry, Kindle Fire, and more. More information on `http://www.sencha.com/`.

Corona is unmatched in giving mobile app developers the ability to develop high quality content at record speeds. This cross-platform tool boosts the team productivity for all major platforms and devices. Thanks to the elegant APIs, tasks like animating objects, creating UI widgets or enabling physics take only a few lines of code. Changes are instantly viewable in the Corona Simulator. More information on `http://www.coronalabs.com/products/corona-sdk/`.

NOTE: The following tools are still available on download platforms (last research conducted in August 2012).

Appendix B

Steps to Protect Your Mobile Phone

Protect personal data when developing an application does not reduce any risk of an outside attack. Users have the feeling that there is no risk when surfing on internet via their smartphone. However, a flaw can be found quickly by malicious people if they do not take necessary measures to prevent any threat. Best practices can reduce consequences of an attack [40]:

- *"**When choosing a mobile phone, consider its security features.** Ask the service provider if the device offers file encryption, the ability for the provider to find and wipe the device remotely, the ability to delete known malicious apps remotely, and authentication features such as device access passwords. If you back up your phone data to a PC, look for an option to encrypt the backup. If you plan to use the device for VPN[1] access, as some users do to access work networks, ask the provider if the device supports certificate-based authentication.*

- ***Configure the device to be more secure.** Many smartphones have a password feature that locks the device until the correct PIN[2] or password is entered. Enable this feature, and choose a reasonably complex password. Enable encryption, remote wipe capabilities, and antivirus software if available.*

- ***Configure web accounts to use secure connections.** Accounts for certain websites can be configured to use secure, encrypted connections (look for HTTPS or SSL in account options pages). Enabling this feature deters attackers from eavesdropping on web sessions. Many popular mail and social networking sites include this option.*

- ***Do not follow links sent in suspicious email or text messages.** Such links may lead to malicious websites.*

- ***Limit exposure of your mobile phone number.** Think carefully before posting your mobile phone number to a public website. Attackers can use software to collect mobile phone numbers from the web and then use those numbers to target attacks.*

[1]VPN : Virtual Private Network
[2]PIN : Personal Identification Number

- ***Carefully consider what information you want stored on the device.*** *Remember that with enough time, sophistication, and access to the device, any attacker could obtain your stored information.*

- ***Be choosy when selecting and installing apps.*** *Do a little research on apps before installing them. Check what permissions the app requires. If the permissions seem beyond what the app should require, do not install the app; it could be a Trojan horse, carrying malicious code in an attractive package.*

- ***Maintain physical control of the device, especially in public or semi-public places.*** *The portability of mobile phones makes them easy to lose or steal.*

- ***Disable interfaces that are not currently in use, such as Bluetooth, infrared, or WiFi.*** *Attackers can exploit vulnerabilities in software that use these interfaces.*

- ***Set Bluetooth-enabled devices to non-discoverable.*** *When in discoverable mode, your Bluetooth-enabled devices are visible to other nearby devices, which may alert an attacker or infected device to target you. When in non-discoverable mode, your Bluetooth-enabled devices are invisible to other unauthenticated devices.*

- ***Avoid joining unknown Wi-Fi networks and using public Wi-Fi hotspots.*** *Attackers can create phony Wi-Fi hotspots designed to attack mobile phones and may patrol public Wi-Fi networks for unsecured devices. Also, enable encryption on your home Wi-Fi network.*

- ***Delete all information stored in a device prior to discarding it.*** *Check the website of the device's manufacturer for information about securely deleting data. Your mobile phone provider may also have useful information on securely wiping your device.*

- ***Be careful when using social networking applications.*** *These apps may reveal more personal information than intended, and to unintended parties. Be especially careful when using services that track your location.*

- ***Do not*** `root` ***or*** `jailbreak` ***the device.*** *Third-party device firmware, which is sometimes used to get access to device features that are locked by default, can contain malicious code or unintentional security vulnerabilities. Altering the firmware could also prevent the device from receiving future operating system updates, which often contain valuable security updates and other feature upgrades."*

In addition, the mobile user must act quickly when his device is stolen. Making some actions such as report the loss or theft to mobile service provider and local authorities. To deter malicious use of device and minimize fraudulent charges. Don't forget to change account credentials by contacting services provider to revoke issued certificates or logging into websites to change password. Finally, if necessary, wipe the phone to delete all data on this one.